

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
3 June 2004 (03.06.2004)

PCT

(10) International Publication Number
WO 2004/046340 A2

- (51) International Patent Classification⁷: C12N
- (21) International Application Number: PCT/US2003/037047
- (22) International Filing Date: 14 November 2003 (14.11.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 60/426,095 14 November 2002 (14.11.2002) US
- (71) Applicant (for all designated States except US): GENENTECH, INC. [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): KRUMMEN, Lynne [US/US]; 3030 22nd Avenue, San Francisco, CA 94132 (US). SHEN, Amy [US/US]; 1828 Parrott Drive, San Mateo, CA 94402 (US).
- (74) Agents: STEFFES, David et al.; Sidley Austin Brown & Wood LLP, 1501 K Street, N.W., Washington, D.C. 20005 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Published:
— without international search report and to be republished upon receipt of that report
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 2004/046340 A2

(54) Title: INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-EXPRESSING PRODUCTION CELL LINES

(57) Abstract: This invention relates to a DNA construct, methods of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest. In one method, stable clones capable of producing a high level of a product of interest are generated from one step of a direct selection immediately after transfection.

INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-
EXPRESSING PRODUCTION CELL LINES

This application claims priority under 35 U.S.C. § 119(e) from U.S. provisional application serial no. 60/426,095, filed November 14, 2002, which is herein incorporated by reference.

BACKGROUND OF THE INVENTION

Field of the Invention

This invention relates to a DNA construct, a method of selecting for high-expressing host cells, a method of producing a protein of interest in high yields and a method of producing eukaryotic cells having multiple copies of a sequence encoding a protein of interest.

Description of Background and Related Art

The discovery of methods for introducing DNA into living host cells in a functional form has provided the key to understanding many fundamental biological processes, and has made possible the production of important proteins and other molecules in commercially useful quantities.

Despite the general success of such gene transfer methods, several common problems exist that may limit the efficiency with which a gene encoding a desired protein can be introduced into and expressed in a host cell. One problem is knowing when the gene has been successfully transferred into recipient cells. A second problem is distinguishing between those cells that contain the gene and those that have survived the transfer procedures but do not contain the gene. A third problem is identifying and isolating those cells that contain the gene and that are expressing high levels of the protein encoded by the gene.

In general, the known methods for introducing genes into eukaryotic cells tend to be highly inefficient. Of the cells in a given culture, only a small proportion take up and express exogenously added DNA, and an even smaller proportion stably maintain that DNA.

Identification of those cells that have incorporated a product gene encoding a desired protein typically is achieved by introducing into the same cells another gene, commonly referred to

as a selectable gene, that encodes a selectable marker. A selectable marker is a protein that is necessary for the growth or survival of a host cell under the particular culture conditions chosen, such as an enzyme that confers resistance to an antibiotic or other drug, or an enzyme that compensates for a metabolic or catabolic defect in the host cell. For example, selectable genes commonly used with eukaryotic cells include the genes for aminoglycoside phosphotransferase (APH), hygromycin phosphotransferase (hyg), dihydrofolate reductase (DHFR), thymidine kinase (tk), neomycin resistance, puromycin resistance, glutamine synthetase, and asparagine synthetase.

The method of identifying a host cell that has incorporated one gene on the basis of expression by the host cell of a second incorporated gene encoding a selectable marker is referred to as cotransfection (or cotransfection). In that method, a gene encoding a desired polypeptide and a selection gene typically are introduced into the host cell simultaneously. In this case of simultaneous cotransfection, the gene encoding the desired polypeptide and the selectable gene may be present on a single DNA molecule or on separate DNA molecules prior to being introduced into the host cells. Wigler *et al.*, Cell, 16:777 (1979). Cells that have incorporated the gene encoding the desired polypeptide then are identified or isolated by culturing the cells under conditions that preferentially allow for the growth or survival of those cells that synthesize the selectable marker encoded by the selectable gene.

The level of expression of a gene introduced into a eukaryotic host cell depends on multiple factors, including gene copy number, efficiency of transcription, messenger RNA (mRNA) processing, stability, and translation efficiency. Accordingly, high level expression of a desired polypeptide typically will involve optimizing one or more of those factors.

For example, the level of protein production may be increased by covalently joining the coding sequence of the gene to a "strong" promoter or enhancer that will give high levels of transcription. Promoters and enhancers are nucleotide sequences that interact specifically with proteins in a host cell that are involved in transcription. Kriegler, Meth. Enzymol., 185:512 (1990); Maniatis *et al.*, Science, 236:1237 (1987). Promoters are located upstream of the coding sequence of a gene and facilitate transcription of the gene by RNA polymerase. Among the eukaryotic promoters that have been identified as strong promoters for high-level expression are the SV40 early promoter, adenovirus major late promoter, mouse metallothionein-I promoter, Rous sarcoma virus long terminal repeat, and human cytomegalovirus immediate early promoter (CMV).

Enhancers stimulate transcription from a linked promoter. Unlike promoters, enhancers are active when placed downstream from the transcription initiation site or at considerable distances from the promoter, although in practice enhancers may overlap physically and functionally with promoters. For example, all of the strong promoters listed above also contain strong enhancers. Bendig, Genetic Engineering, 7:91 (Academic Press, 1988).

The level of protein production also may be increased by increasing the gene copy number in the host cell. One method for obtaining high gene copy number is to directly introduce into the host cell multiple copies of the gene, for example, by using a large molar excess of the product gene relative to the selectable gene during cotransfection. Kaufman, Meth. Enzymol., 185:537 (1990). With this method, however, only a small proportion of the cotransfected cells will contain the product gene at high copy number. Furthermore, because no generally applicable, convenient method exists for distinguishing such cells from the majority of cells that contain fewer copies of the product gene, laborious and time-consuming screening methods typically are required to identify the desired high-copy number transfectants.

Another method for obtaining high gene copy number involves cloning the gene in a vector that is capable of replicating autonomously in the host cell. Examples of such vectors include mammalian expression vectors derived from Epstein-Barr virus or bovine papilloma virus, and yeast 2-micron plasmid vectors. Stephens & Hentschel, Biochem. J., 248:1 (1987); Yates *et al.*, Nature, 313:812 (1985); Beggs, Genetic Engineering, 2:175 (Academic Press, 1981).

Yet another method for obtaining high gene copy number involves gene amplification in the host cell. Gene amplification occurs naturally in eukaryotic cells at a relatively low frequency. Schimke, J. Biol. Chem., 263:5989 (1988). However, gene amplification also may be induced, or at least selected for, by exposing host cells to appropriate selective pressure. For example, in many cases it is possible to introduce a product gene together with an amplifiable gene into a host cell and subsequently select for amplification of the marker gene by exposing the cotransfected cells to sequentially increasing concentrations of a selective agent. Typically the product gene will be coamplified with the marker gene under such conditions.

The most widely used amplifiable gene for that purpose is a DHFR gene, which encodes a dihydrofolate reductase enzyme. The selection conditions used in conjunction with a DHFR gene are the absence of glycine, hypoxanthine and thymidine (GHT) with or without the presence of methotrexate (Mtx). A host cell is cotransfected with a product gene encoding a desired protein

and a DHFR gene, and transfectants are identified by first culturing the cells in GHT-free culture medium that may contain Mtx. A suitable host cell when a wild-type DHFR gene is used is the Chinese Hamster Ovary (CHO) cell line deficient in DHFR activity, prepared and propagated as described by Urlaub & Chasin, Proc. Nat. Acad. Sci. USA, 77:4216 (1980). The transfected cells then are exposed to successively higher amounts of Mtx. This leads to the synthesis of multiple copies of the DHFR gene, and concomitantly, multiple copies of the product gene. Schimke, J. Biol. Chem., 263:5989 (1988); Axel *et al.*, U.S. Patent No. 4,399,216; Axel *et al.*, U.S. Patent No. 4,634,665. Other references directed to co-transfection of a gene together with a genetic marker that allows for selection and subsequent amplification include Kaufman in Genetic Engineering, ed. J. Setlow (Plenum Press, New York), Vol. 9 (1987); Kaufman and Sharp, J. Mol. Biol., 159:601 (1982); Ringold *et al.*, J. Mol. Appl. Genet., 1:165-175 (1981); Kaufman *et al.*, Mol. Cell Biol., 5:1750-1759 (1985); Kaetzel and Nilson, J. Biol. Chem., 263:6244-6251 (1988); Hung *et al.*, Proc. Natl. Acad. Sci. USA, 83:261-264 (1986); Kaufman *et al.*, EMBO J., 6:87-93 (1987); Johnston and Kucey, Science, 242:1551-1554 (1988); Urlaub *et al.*, Cell, 33:405-412 (1983).

To extend the DHFR amplification method to other cell types, a mutant DHFR gene that encodes a protein with reduced sensitivity to methotrexate may be used in conjunction with host cells that contain normal numbers of an endogenous wild-type DHFR gene. Simonsen and Levinson, Proc. Natl. Acad. Sci. USA, 80:2495 (1983); Wigler *et al.*, Proc. Natl. Acad. Sci. USA, 77:3567-3570 (1980); Haber and Schimke, Somatic Cell Genetics, 8:499-508 (1982).

Alternatively, host cells may be co-transfected with the product gene, a DHFR gene, and a dominant selectable gene, such as a neo^r gene. Kim and Wold, Cell, 42:129 (1985); Capon *et al.*, U.S. Pat. No. 4,965,199. Transfectants are identified by first culturing the cells in culture medium containing neomycin (or the related drug G418), and the transfectants so identified then are selected for amplification of the DHFR gene and the product gene by exposure to successively increasing amounts of Mtx.

As will be appreciated from this discussion, the selection of recombinant host cells that express high levels of a desired protein generally is a multi-step process. In the first step, initial transfectants are selected that have incorporated the product gene and the selectable gene. In subsequent steps, the initial transfectants are subject to further selection for high-level expression of the selectable gene and then random screening for high-level expression of the product gene. To identify cells expressing high levels of the desired protein, typically one must screen large numbers

of transfectants. The majority of transfectants produce less than maximal levels of the desired protein. Further, Mtx resistance in DHFR transformants is at least partially conferred by varying degrees of gene amplification. Schimke, Cell, 37:705-713 (1984). The inadequacies of co-expression of the non-selected gene have been reported by Wold *et al.*, Proc. Natl. Acad. Sci. USA, 76:5684-5688 (1979). Instability of the amplified DNA is reported by Kaufman and Schimke, Mol. Cell Biol., 1:1069-1076 (1981); Haber and Schimke, Cell, 26:355-362 (1981); and Fedespiel *et al.*, J. Biol. Chem., 259:9127-9140 (1984).

Several methods have been described for directly selecting such recombinant host cells in a single step. One strategy involves co-transfecting host cells with a product gene and a DHFR gene, and selecting those cells that express high levels of DHFR by directly culturing in medium containing a high concentration of Mtx. Many of the cells selected in that manner also express the co-transfected product gene at high levels Page and Sydenham, Bio/Technology, 9:64 (1991). This method for single-step selection suffers from certain drawbacks that limit its usefulness. High-expressing cells obtained by direct culturing in medium containing a high level of a selection agent may have poor growth and stability characteristics, thus limiting their usefulness for long-term production processes Page and Snyderman, Bio/Technology, 9:64 (1991). Single-step selection for high-level resistance to Mtx may produce cells with an altered, Mtx-resistant DHFR enzyme, or cells that have altered Mtx transport properties, rather than cells containing amplified genes. Haber *et al.*, J. Biol. Chem., 256:9501 (1981); Assaraf and Schimke, Proc. Natl. Acad. Sci. USA, 84:7154 (1987).

Another method involves the use of polycistronic mRNA expression vectors containing a product gene at the 5' end of the transcribed region and a selectable gene at the 3' end. Because translation of the selectable gene at the 3' end of the polycistronic mRNA is inefficient, such vectors exhibit preferential translation of the product gene and require high levels of polycistronic mRNA to survive selection. Kaufman, Meth. Enzymol., 185:487 (1990); Kaufman, Meth. Enzymol., 185:537 (1990); Kaufman *et al.*, EMBO J., 6:187 (1987). Accordingly, cells expressing high levels of the desired protein product may be obtained in a single step by culturing the initial transfectants in medium containing a selection agent appropriate for use with the particular selectable gene. However, the utility of these vectors is variable because of the unpredictable influence of the upstream product reading frame on selectable marker translation and because the upstream reading frame sometimes becomes deleted during methotrexate amplification (Kaufman

et al., J. Mol. Biol., 159:601-621 (1982); Levinson, Methods in Enzymology, San Diego: Academic Press, Inc. (1990)). Later vectors incorporated an internal translation initiation site derived from members of the picornavirus family which is positioned between the product gene and the selectable gene (Pelletier *et al.*, Nature, 334:320 (1988); Jang *et al.*, J. Virol., 63:1651 (1989)).

A third method for single-step selection involves use of a DNA construct with a selectable gene containing an intron within which is located a gene encoding the protein of interest. See U.S. Patent No. 5,043,270 and Abrams *et al.*, J. Biol. Chem., 264(24): 14016-14021 (1989). In yet another single-step selection method, host cells are co-transfected with an intron-modified selectable gene and a gene encoding the protein of interest. See WO 92/17566, published October 15, 1992. The intron-modified gene is prepared by inserting into the transcribed region of a selectable gene an intron of such length that the intron is correctly spliced from the corresponding mRNA precursor at low efficiency, so that the amount of selectable marker produced from the intron-modified selectable gene is substantially less than that produced from the starting selectable gene. These vectors help to insure the integrity of the integrated DNA construct, but transcriptional linkage is not achieved as selectable gene and the protein gene are driven by separate promoters.

Other mammalian expression vectors that have single transcription units have been described. Retroviral vectors have been constructed (Cepko *et al.*, Cell, 37:1053-1062 (1984)) in which a cDNA is inserted between the endogenous Moloney murine leukemia virus (M-MuLV) splice donor and splice acceptor sites which are followed by a neomycin resistance gene. This vector has been used to express a variety of gene products following retroviral infection of several cell types.

A method for selecting recombinant host cells expressing high levels of a desired protein was previously described by the applicants in Lucas *et al.*, Nucleic Acid Research, 24, No. 9: 1774-1779 and U.S. Patent No. 5,561,053. That method utilizes eukaryotic host cells harboring a DNA construct comprising a selectable gene (preferably an amplifiable gene) and a product gene provided 3' to the selectable gene. The selectable gene is positioned within an intron defined by a splice donor site and a splice acceptor site and the selectable gene and product gene are under the transcriptional control of a single transcriptional regulatory region. The splice donor site is generally an efficient splice donor site and thereby regulates expression of the product gene using the transcriptional regulatory region. The transfected cells are cultured so as to express the gene encoding the product in a selective medium which may contain an amplifying agent for sufficient

time to allow cells having multiple copies of the product gene, or cells with a single (or multiple) copy of the gene in a chromosomal loci with high transcriptional activity to be identified.

Other fusion expression constructs have been developed. For example, a fusion of green fluorescent protein with the Zeocin-resistance marker construct has been created. Bennet, R.P. *et al.*, Biotechniques, 24(3):478-82, 1998 March. Such constructs were used to allow visual screening and drug selection of transfected eukaryotic cells.

In another example, human prothrombin was overexpressed in transformed eukaryotic cells using a dominant bifunctional selection and amplification marker. Herlitschka, Sabine E. *et al.*, Protein Expression and Purification, 8, 358-364, 1996 July. In this reference the marker consisted of the murine wild-type dihydrofolate reductase cDNA and the *E. coli* hygromycin phosphotransferase gene fused in frame. The gene of interest is connected, upstream, by the EMCV untranslated region to the fusion marker gene, forming a dicistronic transcription unit.

With the state of the art in mind, it is one object of the present invention to increase the level of homogeneity with regard to expression levels of stable clones transfected with a product gene of interest, by expressing fused selectable markers (i.e. DHFR and puromycin) and a protein of interest from a single promoter.

It is another object to provide a method for selecting stable, recombinant host cells that express high levels of a desired protein product, which method is rapid and convenient to perform, and reduces the numbers of transfected cells which need to be screened. Furthermore, it is an object to allow high levels of single and multiple unit polypeptides to be rapidly generated from clones or pools of stable host cell transfectants.

It is an additional object to provide expression vectors which bias for active integration events (i.e. have an increased tendency to generate transformants wherein the DNA construct is inserted into a region of the genome of the host cell which results in high level expression of the product gene) and can accommodate a variety of product genes without the need for modification.

SUMMARY OF THE INVENTION

Accordingly, the present invention is directed to a DNA construct (DNA molecule) comprising a 5' transcriptional initiation site and a 3' transcriptional termination site, two selectable genes that have been fused into one open reading frame (preferably amplifiable genes) and a product gene provided 3' to the fused selectable genes, a transcriptional regulatory region

regulating transcription of both the fused selectable genes and the product gene, the fused selectable genes positioned within an intron defined by a splice donor site and a splice acceptor site. The splice donor site preferably comprises an effective splice donor sequence as herein defined and thereby regulates expression of the product gene using the transcriptional regulatory region.

In another embodiment, the invention provides a method for producing a product of interest comprising culturing a eukaryotic cell which has been transfected with the DNA construct described above, so as to express the product gene and recovering the product.

In a further embodiment, the invention provides a method for producing eukaryotic cells having multiple copies of the product gene comprising transfecting eukaryotic cells with the DNA construct described above (where the selectable fused genes are amplifiable genes), growing the cells in a selective medium comprising an amplifying agent(s) for a sufficient time for amplification to occur, and selecting cells having multiple copies of the product gene. After transfection of the host cells, most of the transfectants fail to exhibit the selectable phenotype characteristic of the protein encoded by either of the selectable genes, but surprisingly a small proportion of the transfectants do exhibit one or both of the selectable phenotype, and among those transfectants, the majority are found to express high levels of the desired product encoded by the product gene. Thus, the invention provides an improved method for the selection of recombinant host cells expressing high levels of a desired product, which method is useful with a wide variety of eukaryotic host cells and avoids the problems inherent in, and improves upon, existing cell selection technology.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates schematically the construction of the pSV.IPD. The gene for the protein of interest would be inserted at the polylinker site.

Figures 2-1 to 2-4 depict the nucleotide sequence of the pSV.IPUR plasmid used in constructing pSV.IPD (SEQ ID NO 1).

Figures 3-1 to 3-4 depict the nucleotide sequence of the pSV.ID plasmid used in constructing pSV.IPD (SEQ ID NO 2).

Figures 4-1 to 4-4 depict the nucleotide sequence of the pSV.IPD (SEQ ID NO 3).

Figure 5 illustrates schematically the plasmid, pSV.ID.VEGF, used as a control in Example 1.

Figure 6 illustrates schematically the plasmid, pSV.IPD.2C4, used in Example 1 (SEQ ID NO 4).

Figures 7-1 to 7-8 depict the nucleotide sequence of the pSV.IPD.2C4 plasmid used in Example 1.

Figure 8 depicts a FACS analysis of transiently transfected CHO cells with a GTP plasmid in 250ml spinner transfection. FACS analysis was performed 24 hours after transfection.

Figure 9 depicts the expression level of clones from traditional 10nM MTX selection. Cells were transfected with commercial transfection reagent and directly selected in 10 nM MTX. Individual clones were grown in a 96-well plate. Product accumulated for 6 days prior to ELISA.

Figures 10-1 and 10-2 depict the expression level of clones from 25 and 50 nM MTX direct selections, respectively, of SV40-based constructs derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 11 depicts the expression level of clones from 25 nM MTX direct selection of CMV-based construct derived from spinner transfection. The assay was performed the same as in Figure 9.

Figure 12 depicts the titer evaluation in Miniform. Samples were collected every day and submitted to an HPLC protein A assay for titer.

Figure 13-1 to 13-7 depict the nucleotide sequence of the pCMV.IPD.Heterologous polypeptide (HP) plasmid used in Example 3.

Figure 14-1 to 14-8 depicts the nucleotide sequence of the pSV40.IPD.HP plasmid used in Example 3.

Figure 15 illustrates schematically the plasmid, pCMV.IPD.HP, used in Example 3.

Figure 16 illustrates a time line and titer comparison between a traditional selection and direct selection method described in Example 3. Equivalent titers are indicated horizontally across the illustration. For example, the titers for a 200/300nM SV40-plasmid traditional selection, 100nM SV40-plasmid direct selection and 25nm CMV-plasmid direct selection are roughly equivalent.

DESCRIPTION OF THE PREFERRED EMBODIMENT

Definitions:

The "DNA construct" disclosed herein comprises a non-naturally occurring DNA molecule or chemical analog which can either be provided as an isolate or integrated in another DNA molecule *e.g.* in an expression vector or the chromosome of an eukaryotic host cell.

The term "selectable gene" as used herein refers to a DNA that encodes a selectable marker necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. Accordingly, a host cell that is transformed with a selectable gene will be capable of growth or survival under certain cell culture conditions wherein a non-transfected host cell is not capable of growth or survival. Typically, a selectable gene will confer resistance to a drug or compensate for a metabolic or catabolic defect in the host cell. Examples of selectable genes are provided in the following table. See also Kaufman, Methods in Enzymology, 185: 537-566 (1990), for a review of these.

"Fused selectable genes" as used herein refers to a DNA that encodes at least two selectable markers in the same open reading frame and inserted into an intron sequence.

TABLE 1**Examples of Selectable Genes and their Selection Agents**

Selection Agent	Selectable Gene
Puromycin	Puromycin-N-acetyltransferase
Methotrexate	Dihydrofolate reductase
Cadmium	Metallothionein
PALA	CAD
Xyl-A-or adenosine and 2'-deoxycycoformycin	Adenosine deaminase
Adenine, azaserine, and coformycin	Adenylate deaminase
6-Azaauridine, pyrazofuran	UMP Synthetase
Mycophenolic acid	IMP 5'-dehydrogenase
Mycophenolic acid with limiting xanthine	Xanthine-guanine phosphoribosyltransferase
Hypoxanthine, aminopterin, and thymidine (HAT)	Mutant HGPRTase or mutant thymidine kinase
5-Fluorodeoxyuridine	Thymidylate synthetase
Multiple drugs e.g. adriamycin, vincristine or colchicine	P-glycoprotein 170
Aphidicolin	Ribonucleotide reductase
Methionine sulfoximine	Glutamine synthetase

β -Aspartyl hydroxamate or Albizziin	Asparagine synthetase
Canavanine	Arginosuccinate synthetase
α -Difluoromethylornithine	Ornithine decarboxylase
Compactin	HMG-CoA reductase
Tunicamycin	N-Acetylglucosaminyl transferase
Borrelidin	Threonyl-tRNA synthetase
Ouabain	Na ⁺ K ⁺ -ATPase

The preferred selectable genes are amplifiable genes. As used herein, the term "amplifiable gene" refers to a gene which is amplified (*i.e.* additional copies of the gene are generated which survive in intrachromosomal or extrachromosomal form) under certain conditions. The amplifiable gene(s) usually encodes an enzyme (*i.e.* an amplifiable marker) which is required for growth of eukaryotic cells under those conditions. For example, the gene may encode DHFR which is amplified when a host cell transformed therewith is grown in Mtx. According to Kaufman, the selectable genes in Table 1 above can also be considered amplifiable genes. An example of a selectable gene which is generally not considered to be an amplifiable gene is the neomycin resistance gene (Cepko *et al.*, *supra*).

As used herein, "selective medium" refers to nutrient solution used for growing eukaryotic cells which have the selectable gene(s) and therefore is deficient in components supplied by the selectable gene or includes a "selection agent". Commercially available media based on formulations such as Ham's F10 (Sigma), Minimal Essential Medium ((MEM), Sigma), RPMI-1640 (Sigma), and Dulbecco's Modified Eagle's Medium ((DMEM), Sigma) are exemplary nutrient solutions. In addition, any of the media described in Ham and Wallace, Meth. Enz., 58:44 (1979), Barnes and Sato, Anal. Biochem., 102:255 (1980), U.S. Patent Nos. 4,767,704; 4,657,866; 4,927,762; or 4,560,655; WO 90/03430; WO 87/00195; U.S. Patent Re. 30,985; or U.S. Patent No. 5,122,469, the disclosures of all of which are incorporated herein by reference, may be used as culture media. Any of these media may be supplemented as necessary with hormones and/or other

growth factors (such as insulin, transferrin, or epidermal growth factor), salts (such as sodium chloride, calcium, magnesium, and phosphate), buffers (such as HEPES), nucleosides (such as adenosine and thymidine), antibiotics (such as GentamycinTM drug), trace elements (defined as inorganic compounds usually present at final concentrations in the micromolar range), and glucose or an equivalent energy source. Any other necessary supplements may also be included at appropriate concentrations that would be known to those skilled in the art. The preferred nutrient solution comprises fetal bovine serum.

The term "selection agent" refers to a substance that interferes with the growth or survival of a host cell possibly because the cell is deficient in a particular selectable gene. Examples of selection agents are presented in Table 1 above. The selection agent preferably comprises an "amplifying agent" which is defined for purposes herein as an agent for amplifying copies of the amplifiable gene or causing integration of multiple copies of the amplifiable gene into the genome, such as Mtx if the amplifiable gene is DHFR. See Table 1 for examples of amplifying agents.

As used herein, the terms "direct selection" or "direct culturing" means the first exposure to selective conditions either without MTX or GHT or with MTX, and production of a heterologous polypeptide in an amount of about 250mg/l, 400mg/l, 600mg/l or 800mg/l up to about 1g/l or more.

As used herein, the term "transcriptional initiation site" refers to the nucleic acid in the DNA construct corresponding to the first nucleic acid incorporated into the primary transcript, *i.e.*, the mRNA precursor, which site is generally provided at, or adjacent to, the 5' end of the DNA construct.

The term "transcriptional termination site" refers to a sequence of DNA, normally represented at the 3' end of the DNA construct, that causes RNA polymerase to terminate transcription.

As used herein, "transcriptional regulatory region" refers to a region of the DNA construct that regulates transcription of the selectable gene and the product gene. The transcriptional regulatory region normally refers to a promoter sequence (*i.e.* a region of DNA involved in binding of RNA polymerase to initiate transcription) which can be constitutive or inducible and, optionally, an enhancer (*i.e.* a *cis*-acting DNA element, usually from about 10-300 bp, that acts on a promoter to increase its transcription).

As used herein, "product gene" refers to DNA that encodes a desired protein or polypeptide product. Any product gene that is capable of expression in a host cell may be used, although the methods of the invention are particularly suited for obtaining high-level expression of a product gene that is not also a selectable or amplifiable gene. Accordingly, the protein or polypeptide encoded by a product gene typically will be one that is not necessary for the growth or survival of a host cell under the particular cell culture conditions chosen. For example, product genes suitably encode a peptide, or may encode a polypeptide sequence of amino acids for which the chain length is sufficient to produce higher levels of tertiary and/or quaternary structure.

Examples of bacterial polypeptides or proteins include, *e.g.*, alkaline phosphatase and β -lactamase. Examples of mammalian polypeptides or proteins include molecules such as renin; a growth hormone, including human growth hormone, and bovine growth hormone; growth hormone releasing factor; parathyroid hormone; thyroid stimulating hormone; lipoproteins; alpha-1-antitrypsin; insulin A-chain; insulin B-chain; proinsulin; follicle stimulating hormone; calcitonin; luteinizing hormone; glucagon; clotting factors such as factor VIIIc, factor IX, tissue factor, and von Willebrand's factor; anti-clotting factors such as Protein C; atrial natriuretic factor; lung surfactant; a plasminogen activator, such as urokinase or human urine or tissue-type plasminogen activator (t-PA); bombesin; thrombin; hemopoietic growth factor; tumor necrosis factor-alpha and -beta; enkephalinase; RANTES (regulated on activation normally T-cell expressed and secreted); human macrophage inflammatory protein (MIP-1-alpha); a serum albumin such as human serum albumin; mullerian-inhibiting substance; relaxin A-chain; relaxin B-chain; prorelaxin; mouse gonadotropin-associated peptide; a microbial protein, such as beta-lactamase; DNase; inhibin; activin; vascular endothelial growth factor (VEGF); receptors for hormones or growth factors; integrin; protein A or D; rheumatoid factors; a neurotrophic factor such as bone-derived neurotrophic factor (BDNF), neurotrophin-3, -4, -5, or -6 (NT-3, NT-4, NT-5, or NT-6), or a nerve growth factor such as NGF- β ; platelet-derived growth factor (PDGF); fibroblast growth factor such as aFGF and bFGF; epidermal growth factor (EGF); transforming growth factor (TGF) such as TGF-alpha and TGF-beta, including TGF- β 1, TGF- β 2, TGF- β 3, TGF- β 4, or TGF- β 5; insulin-like growth factor-I and -II (IGF-I and IGF-II); des(1-3)-IGF-I (brain IGF-I), insulin-like growth factor binding proteins; CD proteins such as CD-3, CD-4, CD-8, and CD-19; erythropoietin; osteoinductive factors; immunotoxins; a bone morphogenetic protein (BMP); an interferon such as interferon-alpha, -beta, and -gamma; colony stimulating factors (CSFs), *e.g.*, M-CSF, GM-CSF,

and G-CSF; interleukins (ILs), e.g., IL-1 to IL-10; superoxide dismutase; T-cell receptors; surface membrane proteins; decay accelerating factor; viral antigen such as, for example, a portion of the AIDS envelope; transport proteins; homing receptors; addressins; regulatory proteins; antibodies; chimeric proteins such as immunoadhesins and fragments of any of the above-listed polypeptides.

The product gene preferably does not consist of an anti-sense sequence for inhibiting the expression of a gene present in the host. Preferred proteins herein are therapeutic proteins such as TGF- β , TGF- α , PDGF, EGF, FGF, IGF-I, DNase, plasminogen activators such as t-PA, clotting factors such as tissue factor and factor VIII, hormones such as relaxin and insulin, cytokines such as IFN- γ , chimeric proteins such as TNF receptor IgG immunoadhesin (TNFr-IgG) or antibodies such as anti-IgE. An example of an antibody that can be produced with the pSV.IDP plasmid (Figure 4) is anti-HER2 Neu antibody, 2C4, as provided in Example 1, *supra*.

The term "intron" as used herein refers to a nucleotide sequence present within the transcribed region of a gene or within a messenger RNA precursor, which nucleotide sequence is capable of being excised, or spliced, from the messenger RNA precursor by a host cell prior to translation. Introns suitable for use in the present invention are suitably prepared by any of several methods that are well known in the art, such as purification from a naturally occurring nucleic acid or *de novo* synthesis. The introns present in many naturally occurring eukaryotic genes have been identified and characterized. Mount, Nuc. Acids Res., **10**:459 (1982). Artificial introns comprising functional splice sites also have been described. Winey *et al.*, Mol. Cell Biol., **9**:329 (1989); Gattermann *et al.*, Mol. Cell Biol., **9**:1526 (1989). Introns may be obtained from naturally occurring nucleic acids, for example, by digestion of a naturally occurring nucleic acid with a suitable restriction endonuclease, or by PCR cloning using primers complementary to sequences at the 5' and 3' ends of the intron. Alternatively, introns of defined sequence and length may be prepared synthetically using various methods in organic chemistry. Narang *et al.*, Meth. Enzymol., **68**:90 (1979); Caruthers *et al.*, Meth. Enzymol., **154**:287 (1985); Froehler *et al.*, Nuc. Acids Res., **14**:5399 (1986).

As used herein "splice donor site" or "SD" refers to the DNA sequence immediately surrounding the exon-intron boundary at the 5' end of the intron, where the "exon" comprises the nucleic acid 5' to the intron. Many splice donor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., **195**:247-259 (1987) provides a review of these. An "efficient splice donor sequence" refers to a nucleic acid sequence encoding a splice donor site wherein the efficiency of splicing of

messenger RNA precursors having the splice donor sequence is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. Examples of efficient splice donor sequences include the wild type (WT) ras splice donor sequence and the GAC:GTAAGT sequence of Example 3. Other efficient splice donor sequences can be readily selected using the techniques for measuring the efficiency of splicing disclosed herein.

The terms "PCR" and "polymerase chain reaction" as used herein refer to the *in vitro* amplification method described in US Patent No. 4,683,195 (issued July 28, 1987). In general, the PCR method involves repeated cycles of primer extension synthesis, using two DNA primers capable of hybridizing preferentially to a template nucleic acid comprising the nucleotide sequence to be amplified. The PCR method can be used to clone specific DNA sequences from total genomic DNA, cDNA transcribed from cellular RNA, viral or plasmid DNAs. Wang & Mark, in PCR Protocols, pp.70-75 (Academic Press, 1990); Scharf, in PCR Protocols, pp. 84-98; Kawasaki & Wang, in PCR Technology, pp. 89-97 (Stockton Press, 1989). Reverse transcription-polymerase chain reaction (RT-PCR) can be used to analyze RNA samples containing mixtures of spliced and unspliced mRNA transcripts. Fluorescently tagged primers designed to span the intron are used to amplify both spliced and unspliced targets. The resultant amplification products are then separated by gel electrophoresis and quantitated by measuring the fluorescent emission of the appropriate band(s). A comparison is made to determine the amount of spliced and unspliced transcripts present in the RNA sample.

One preferred splice donor sequence is a "consensus splice donor sequence". The nucleotide sequences surrounding intron splice sites, which sequences are evolutionarily highly conserved, are referred to as "consensus splice donor sequences". In the mRNAs of higher eukaryotes, the 5' splice site occurs within the consensus sequence AG:GUAAGU (wherein the colon denotes the site of cleavage and ligation). In the mRNAs of yeast, the 5' splice site is bounded by the consensus sequence :GUAUGU. Padgett, *et al.*, Ann. Rev. Biochem., 55:1119 (1986).

The expression "splice acceptor site" or "SA" refers to the sequence immediately surrounding the intron-exon boundary at the 3' end of the intron, where the "exon" comprises the nucleic acid 3' to the intron. Many splice acceptor sites have been characterized and Ohshima *et al.*, J. Mol. Biol., 195:247-259 (1987) provides a review of these. The preferred splice acceptor site is an efficient splice acceptor site which refers to a nucleic acid sequence encoding a splice

acceptor site wherein the efficiency of splicing of messenger RNA precursors having the splice acceptor site is between about 80 to 99% and preferably 90 to 95% as determined by quantitative PCR. The splice acceptor site may comprise a consensus sequence. In the mRNAs of higher eukaryotes, the 3' splice acceptor site occurs within the consensus sequence (U/C)₁₁NCAG:G. In the mRNAs of yeast, the 3' acceptor splice site is bounded by the consensus sequence (C/U)AG:. Padgett, *et al.*, *supra*.

As used herein "culturing for sufficient time to allow amplification to occur" refers to the act of physically culturing the eukaryotic host cells which have been transformed with the DNA construct in cell culture media containing the amplifying agent, until the copy number of the amplifiable gene (and preferably also the copy number of the product gene) in the host cells has increased relative to the transformed cells prior to this culturing.

The term "expression" as used herein refers to transcription or translation occurring within a host cell. The level of expression of a product gene in a host cell may be determined on the basis of either the amount of corresponding mRNA that is present in the cell or the amount of the protein encoded by the product gene that is produced by the cell. For example, mRNA transcribed from a product gene is desirably quantitated by northern hybridization or quantitative real-time PCR. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 7.3-7.57 (Cold Spring Harbor Laboratory Press, 1989). Protein encoded by a product gene can be quantitated either by assaying for the biological activity of the protein or by employing assays that are independent of such activity, such as western blotting or radioimmunoassay using antibodies that are capable of reacting with the protein. Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, pp. 18.1-18.88 (Cold Spring Harbor Laboratory Press, 1989).

Modes for Carrying Out the Invention

Methods and compositions are provided for enhancing the stability and/or copy number of a transcribed sequence in order to allow for elevated levels of a RNA sequence of interest. In general, the methods of the present invention involve transfecting a eukaryotic host cell with an expression vector comprising both a product gene encoding a desired polypeptide and fused selectable genes.

Selectable genes and product genes may be obtained from genomic DNA, cDNA transcribed from cellular RNA, or by *in vitro* synthesis. For example, libraries are screened with

probes (such as antibodies or oligonucleotides of about 20-80 bases) designed to identify the selectable gene or the product gene (or the protein(s) encoded thereby). Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures as described in chapters 10-12 of Sambrook *et al.*, Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the selectable gene or product gene is to use PCR methodology as described in section 14 of Sambrook *et al.*, *supra*.

A preferred method of practicing this invention is to use carefully selected oligonucleotide sequences to screen cDNA libraries from various tissues known to contain the selectable gene or product gene. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized.

The oligonucleotide generally is labeled such that it can be detected upon hybridization to DNA in the library being screened. The preferred method of labeling is to use ³²P- labeled ATP with polynucleotide kinase, as is well known in the art, to radiolabel the oligonucleotide. However, other methods may be used to label the oligonucleotide, including, but not limited to, biotinylation or enzyme labeling.

Sometimes, the DNA encoding the fused selectable genes and product gene is preceded by DNA encoding a signal sequence having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the expression vector, or it may be a part of the selectable gene or product gene that is inserted into the expression vector. If a heterologous signal sequence is used, it preferably is one that is recognized and processed (*i.e.*, cleaved by a signal peptidase) by the host cell. For yeast secretion the native signal sequence may be substituted by, *e.g.*, the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Pat. No. 5,010,182 issued 23 April 1991), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression the native signal sequence of the protein of interest is satisfactory, although other mammalian signal sequences may be suitable, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders, for example, the herpes simplex gD signal. The DNA for such precursor region is ligated in reading frame to the fused selectable genes or product gene.

As shown in Figure 1, the fused selectable genes are generally provided at the 5' end of the DNA construct and are followed by the product gene (which would be inserted into the linker site). Therefore, the full-length (non-spliced) message will contain, for example, the PURO-DHFR fusion as the first open reading frame and will therefore generate PURO-DHFR protein to allow selection of stable transfectants. The full length message is not expected to generate appreciable amounts of the protein of interest as the second AUG in a dicistronic message is an inefficient initiator of translation in mammalian cells (Kozak, J. Cell Biol., 115: 887-903 (1991)).

The fused selectable genes are positioned within an intron. Introns are noncoding nucleotide sequences, normally present within many eukaryotic genes, which are removed from newly transcribed mRNA precursors in a multiple-step process collectively referred to as splicing.

A single mechanism is thought to be responsible for the splicing of mRNA precursors in mammalian, plant, and yeast cells. In general, the process of splicing requires that the 5' and 3' ends of the intron be correctly cleaved and the resulting ends of the mRNA be accurately joined, such that a mature mRNA having the proper reading frame for protein synthesis is produced. Analysis of a variety of naturally occurring and synthetically constructed mutant genes has shown that nucleotide changes at many of the positions within the consensus sequences at the 5' and 3' splice sites have the effect of reducing or abolishing the synthesis of mature mRNA. Sharp, Science, 235:766 (1987); Padgett, *et al.*, Ann. Rev. Biochem., 55:1119 (1986); Green, Ann. Rev. Genet., 20:671 (1986). Mutational studies also have shown that RNA secondary structures involving splicing sites can affect the efficiency of splicing. Solnick, Cell, 43:667 (1985); Konarska, *et al.*, Cell, 42:165 (1985).

The length of the intron may also affect the efficiency of splicing. By making deletion mutations of different sizes within the large intron of the rabbit beta-globin gene, Wieringa, *et al.* determined that the minimum intron length necessary for correct splicing is about 69 nucleotides. Cell, 37:915 (1984). Similar studies of the intron of the adenovirus E1A region have shown that an intron length of about 78 nucleotides allows correct splicing to occur, but at reduced efficiency. Increasing the length of the intron to 91 nucleotides restores normal splicing efficiency, whereas truncating the intron to 63 nucleotides abolishes correct splicing. Ulfendahl, *et al.*, Nuc. Acids Res., 13:6299 (1985).

To be useful in the invention, the intron must have a length such that splicing of the intron from the mRNA is efficient. The preparation of introns of differing lengths is a routine matter,

involving methods well known in the art, such as *de novo* synthesis or *in vitro* deletion mutagenesis of an existing intron. Typically, the intron will have a length of at least about 150 nucleotides, since introns which are shorter than this tend to be spliced less efficiently. The upper limit for the length of the intron can be up to 30 kB or more. However, as a general proposition, the intron is generally less than about 10 kB in length.

The intron is modified to contain the fused selectable genes not normally present within the intron using any of the various known methods for modifying a nucleic acid *in vitro*. Typically, the fused selectable genes will be introduced into an intron by first cleaving the intron with a restriction endonuclease, and then covalently joining the resulting restriction fragments to the fused selectable genes in the correct orientation for host cell expression, for example by ligation with a DNA ligase enzyme.

The DNA construct is dicistronic, *i.e.* the fused selectable genes and product gene are both under the transcriptional control of a single transcriptional regulatory region. As mentioned above, the transcriptional regulatory region comprises a promoter. Suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase (Hitzeman *et al.*, J. Biol. Chem., 255:2073 (1980)) or other glycolytic enzymes (Hess *et al.*, J. Adv. Enzyme Reg., 7:149 (1968); and Holland, Biochemistry, 17:4900 (1978)), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in Hitzeman *et al.*, EP 73,657A. Yeast enhancers also are advantageously used with yeast promoters.

Expression control sequences are known for eukaryotes. Virtually all eukaryotic genes have an AT-rich region located approximately 25 to 30 bases upstream from the site where transcription is initiated. Another sequence found 70 to 80 bases upstream from the start of transcription of many genes is a CXCAAT region where X may be any nucleotide.

Product gene transcription from vectors in mammalian host cells is controlled by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, a retrovirus, hepatitis-B virus and most preferably Simian Virus 40 (SV40) or cytomegalovirus (CMV), from heterologous mammalian promoters, e.g. the actin promoter or an immunoglobulin promoter, from heat-shock promoters, and from the promoter normally associated with the product gene, provided such promoters are compatible with the host cell systems. Promoters endogenous to the host cell system, such as the CHO Elongation Factor 1 alpha promoter may also be used.

The early and late promoters of the SV40 virus are conveniently obtained as an SV40 restriction fragment that also contains the SV40 viral origin of replication. Fiers *et al.*, Nature, 273:113 (1978); Mulligan and Berg, Science, 209:1422-1427 (1980); Pavlakis *et al.*, Proc. Natl. Acad. Sci. USA, 78:7398-7402 (1981). The immediate early promoter of the human cytomegalovirus (CMV) is conveniently obtained as a *Hind*III E restriction fragment. Greenaway *et al.*, Gene, 18:355-360 (1982). A system for expressing DNA in mammalian hosts using the bovine papilloma virus as a vector is disclosed in U.S. 4,419,446. A modification of this system is described in U.S. 4,601,978. See also Gray *et al.*, Nature, 295:503-508 (1982) on expressing cDNA encoding immune interferon in monkey cells; , Reyes *et al.*, Nature, 297:598-601 (1982) on expression of human β -interferon cDNA in mouse cells under the control of a thymidine kinase promoter from herpes simplex virus, Canaani and Berg, Proc. Natl. Acad. Sci. USA, 79:5166-5170 (1982) on expression of the human interferon β 1 gene in cultured mouse and rabbit cells, and Gorman *et al.*, Proc. Natl. Acad. Sci. USA, 79:6777-6781 (1982) on expression of bacterial CAT sequences in CV-1 monkey kidney cells, chicken embryo fibroblasts, Chinese hamster ovary cells, HeLa cells, and mouse NIH-3T3 cells using the Rous sarcoma virus long terminal repeat as a promoter.

Preferably the transcriptional regulatory region in higher eukaryotes comprises an enhancer sequence. Enhancers are relatively orientation and position independent having been found 5' (Lainins *et al.*, Proc. Natl. Acad. Sci. USA, 78:993 (1981)) and 3' (Lusky *et al.*, Mol. Cell Bio., 3:1108 (1983)) to the transcription unit, within an intron (Banerji *et al.*, Cell, 33:729 (1983)) as well as within the coding sequence itself (Osborne *et al.*, Mol. Cell Bio., 4:1293 (1984)). Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -

fetoprotein and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer (CMV), the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. See also Yaniv, Nature, 297:17-18 (1982) on enhancing elements for activation of eukaryotic promoters. The enhancer may be spliced into the vector at a position 5' or 3' to the product gene, but is preferably located at a site 5' from the promoter.

The DNA construct of the present invention has a transcriptional initiation site following the transcriptional regulatory region and a transcriptional termination region following the product gene (see, e.g., Figure 1). These sequences are provided in the DNA construct using techniques which are well known in the art.

The DNA construct normally forms part of an expression vector which may have other components such as an origin of replication (*i.e.*, a nucleic acid sequence that enables the vector to replicate in one or more selected host cells) and, if desired, one or more additional selectable gene(s). Construction of suitable vectors containing the desired coding and control sequences employs standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to generate the plasmids required.

Generally, in cloning vectors the origin of replication is one that enables the vector to replicate independently of the host chromosomal DNA, and includes origins of replication or autonomously replicating sequences. Such sequences are well known. The 2 μ plasmid origin of replication is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells. Generally, the origin of replication component is not needed for mammalian expression vectors (the SV40 origin may typically be used only because it contains the early promoter).

Most expression vectors are "shuttle" vectors, *i.e.*, they are capable of replication in at least one class of organisms but can be transfected into another organism for expression. For example, a vector is cloned in *E. coli* and then the same vector is transfected into yeast or mammalian cells for expression even though it is not capable of replicating independently of the host cell chromosome.

For analysis to confirm correct sequences in plasmids constructed, plasmids from the transformants are prepared, analyzed by restriction, and/or sequenced by the method of Messing *et*

al., Nucleic Acids Res., 9:309 (1981) or by the method of Maxam *et al.*, Methods in Enzymology, 65:499 (1980).

The expression vector having the DNA construct prepared as discussed above is transformed into a eukaryotic host cell. Suitable host cells for cloning or expressing the vectors herein are yeast or higher eukaryote cells.

Eukaryotic microbes such as filamentous fungi or yeast are suitable hosts for vectors containing the product gene. *Saccharomyces cerevisiae*, or common baker's yeast, is the most commonly used among lower eukaryotic host microorganisms. However, a number of other genera, species, and strains are commonly available and useful herein, such as *S. pombe* (Beach and Nurse, Nature, 290:140 (1981)), *Kluyveromyces lactis* (Louvencourt *et al.*, J. Bacteriol., 737 (1983)), *Kyarrowia* (EP 402,226), *Pichia pastoris* (EP 183,070), *Trichoderma reesia* (EP 244,234), *Neurospora crassa* (Case *et al.*, Proc. Natl. Acad. Sci. USA, 76:5259-5263 (1979)), and *Aspergillus* hosts such as *A. nidulans* (Ballance *et al.*, Biochem. Biophys. Res. Commun., 112:284-289 (1983); Tilburn *et al.*, Gene, 26:205-221 (1983); Yelton *et al.*, Proc. Natl. Acad. Sci. USA, 81:1470-1474 (1984)) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 (1985)).

Suitable host cells for the expression of the product gene are derived from multicellular organisms. Such host cells are capable of complex processing and glycosylation activities. In principle, any higher eukaryotic cell culture is workable, whether from vertebrate or invertebrate culture. Examples of invertebrate cells include plant and insect cells. Numerous baculoviral strains and variants and corresponding permissive insect host cells from hosts such as *Spodoptera frugiperda* (caterpillar), *Aedes aegypti* (mosquito), *Aedes albopictus* (mosquito), *Drosophila melanogaster* (fruitfly), and *Bombyx mori* host cells have been identified. See, e.g., Luckow *et al.*, Bio/Technology, 6:47-55 (1988); Miller *et al.*, in Genetic Engineering, Setlow, J.K. *et al.*, eds., Vol. 8 (Plenum Publishing, 1986), pp. 277-279; and Maeda *et al.*, Nature, 315:592-594 (1985). A variety of such viral strains are publicly available, e.g., the L-1 variant of *Autographa californica* NPV and the Bm-5 strain of *Bombyx mori* NPV, and such viruses may be used as the virus herein according to the present invention, particularly for transfection of *Spodoptera frugiperda* cells.

Plant cell cultures of cotton, corn, potato, soybean, petunia, tomato, and tobacco can be utilized as hosts. Typically, plant cells are transfected by incubation with certain strains of the bacterium *Agrobacterium tumefaciens*, which has been previously manipulated to contain the product gene. During incubation of the plant cell culture with *A. tumefaciens*, the product gene is

transferred to the plant cell host such that it is transfected, and will, under appropriate conditions, express the product gene. In addition, regulatory and signal sequences compatible with plant cells are available, such as the nopaline synthase promoter and polyadenylation signal sequences. Depicker *et al.*, J. Mol. Appl. Gen., 1:561 (1982). In addition, DNA segments isolated from the upstream region of the T-DNA 780 gene are capable of activating or increasing transcription levels of plant-expressible genes in recombinant DNA-containing plant tissue. EP 321,196 published 21 June 1989.

However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture (tissue culture) has become a routine procedure in recent years (Tissue Culture, Academic Press, Kruse and Patterson, editors (1973)). Examples of useful mammalian host cell lines are monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham *et al.*, J. Gen Virol., 36:59 (1977)); baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); dp12.CHO cells (EP 307,247 published 15 March 1989); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); monkey kidney cells (CV1 ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor (MMT 060562, ATCC CCL51); TRI cells (Mather *et al.*, Annals N.Y. Acad. Sci., 383:44-68 (1982)); MRC 5 cells; FS4 cells; and a human hepatoma line (Hep G2).

Host cells are transformed with the above-described expression or cloning vectors of this invention and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences.

Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw *et al.*, Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) may be used. General aspects of mammalian cell host system transformations have been described by Axel in U.S. 4,399,216 issued 16 August 1983. Transformations into yeast are typically carried out according to the method of Van Solingen *et al.*, J. Bact., 130:946 (1977) and Hsiao *et al.*, Proc. Natl. Acad. Sci. (USA), 76:3829

(1979). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used.

In preferred embodiments the DNA is introduced into the host cells using electroporation, lipofection or polyfection techniques. In a particularly preferred embodiment, the transfection is performed in a spinner vessel as illustrated by Example 3 or in some other form of suspension culture. Transfection performed in a spinner vessel is also referred to as "spinner transfection". Culturing the cells in suspension allows them to reach a cell density of at least about 5×10^5 /ml and more preferably at least about 1.5×10^6 /ml prior to transfection. See Andreason, J. Tiss. Cult. Meth., 15:56-62 (1993), for a review of electroporation techniques useful for practicing the claimed invention. It was discovered that these techniques for introducing the DNA construct into the host cells are preferable over calcium phosphate precipitation techniques insofar as the latter could cause the DNA to break up and form concatemers.

The mammalian host cells used to express the product gene herein may be cultured in a variety of media as discussed in the definitions section above. The media is formulated to provide selective nutrient conditions or a selection agent to select transformed host cells which have taken up the DNA construct (either as an intra- or extra-chromosomal element). To achieve selection of the transformed eukaryotic cells, the host cells may be grown in cell culture plates and individual colonies expressing one or both of the selectable genes (and thus the product gene) can be isolated and grown in growth medium under defined conditions. The host cells are then analyzed for transcription and/or transformation as discussed below. The culture conditions, such as temperature, pH, and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA (Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)), dot blotting (DNA or mRNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Various labels may be employed, most commonly radioisotopes, particularly ^{32}P . However, other techniques may also be employed, such as using biotin-modified nucleotides for introduction into a polynucleotide. The biotin then serves as the site for binding to avidin or antibodies, which may be labeled with a wide variety of labels, such as radionuclides, fluorescence, enzymes, or the like. Alternatively, antibodies may be employed that can recognize specific

duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. With immunohistochemical staining techniques, a cell sample is prepared, typically by dehydration and fixation, followed by reaction with labeled antibodies specific for the gene product coupled, where the labels are usually visually detectable, such as enzymatic labels, fluorescent labels, luminescent labels, and the like. A particularly sensitive staining technique suitable for use in the present invention is described by Hsu *et al.*, Am. J. Clin. Path., 75:734-738 (1980).

In the preferred embodiment protein expression is measured using ELISA as described in Example 1 herein.

The product of interest preferably is recovered from the culture medium as a secreted polypeptide, although it also may be recovered from host cell lysates when directly expressed without a secretory signal. When the product gene is expressed in a recombinant cell other than one of human origin, the product of interest is completely free of proteins or polypeptides of human origin. However, it is necessary to purify the product of interest from recombinant cell proteins or polypeptides to obtain preparations that are substantially homogeneous as to the product of interest. As a first step, the culture medium or lysate is centrifuged to remove particulate cell debris. The product of interest thereafter is purified from contaminant soluble proteins and polypeptides, for example, by fractionation on immunoaffinity or ion-exchange columns; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel electrophoresis using, for example, Sephadex G-75; chromatography on plasminogen columns to bind the product of interest and protein A Sepharose columns to remove contaminants such as IgG.

The following examples are offered by way of illustration only and are not intended to limit the invention in any manner. All patent and literature references cited herein are expressly incorporated by reference.

EXAMPLE 1

2C4 production using the fusion construct expression vector

Vectors related to those described by Lucas et al (Lucas BK, Giere LM, DeMarco RA, Shen A, Chisholm V and Crowley C. High-level production of recombinant proteins in CHO cells using a dicistronic DHFR intron expression vector. (1996) *Nucleic Acids Res.* 24(9), 1774-1779.), which contain an intron between the SV40 promoter and enhancer and the cDNA that encodes the polypeptide of interest, were constructed. The intron is bordered on its 3' and 5' ends, respectively, by a splice donor site derived from cytomegalovirus immediate early gene (CMVIE), and a splice acceptor site from an IgG heavy chain variable region (V_H) gene (Eaton *et al.*, *Biochem.*, 25:8343 (1986)). The splice sites selected provide slightly inefficient splicing such that only about 90% of the transcripts produced are intron free. Previous studies have demonstrated that when a selectable marker such as DHFR is integrated within this intron, as in the plasmid pSV.ID, marker gene transcription proceeds from any unspliced transcripts, providing a highly efficient means of maintaining linkage between the expression of the marker gene and the cDNA of interest as well as enhanced product expression relative to expression of the marker gene.

Vectors containing a murine puromycin/DHFR fusion sequence in the intron following the SV40 promoter elements were constructed by linearizing a pSV.IPUR plasmid, which contained the puromycin resistance gene in an intron following the SV40 promoter/enhancer (pSV.IPUR, Figures 1 and 2), with Hpa I immediately following the end of the puromycin ORF. A 564 bp PCR fragment containing the entire coding region for the murine DHFR gene was subsequently ligated into this linearized vector 3' of the puromycin resistance gene. The stop codon TAG between the puromycin resistance gene and the DHFR gene was deleted by site-directed mutagenesis resulting in a pSV.I plasmid containing a Puro/DHFR fusion gene within the intron of the expression cassette (pSV.IPD, Figures 1 and 4).

The cDNA of the Heavy chain (HC) and light chain (LC) sequences of an anti-HER2 Neu antibody, 2C4, were inserted into pSV.IPD as shown in Figure 6. The sequence of the resulting pSV.IPD.2C4 vector is shown in Figure 7. Data collected using the pSV.IPD.2C4 vector are shown in Table 2.

Additionally, a vector containing only a murine DHFR sequence within the intron (pSV.ID) was prepared. The DNA sequence for the pSV.ID vector is shown in Figure 3. The preparation of such vectors is disclosed in U.S. Patent No. 5,561,053, which is herein incorporated by reference. Into that vector, the HC and LC sequences of monoclonal antibodies to VEGF were inserted. The sequence of the resulting pSV.ID.VEGF vector is shown in Figure 5.

Plasmid DNA's that contained either the Puro/DHFR fusion sequences in the intron or murine DHFR alone preceding cDNA sequences for HC and LC of 2C4 and anti-VEGF, respectively were introduced into CHO DHFR minus cells by lipofection. Briefly, for transfection, 4 million CHO DUX-B11 (DHFR minus) were seeded in 10 cm plates the day before transfection. On the day of transfection, 4 ug DNA was mixed with 300 ul of serum free medium and 25 ul of polyfect from Qiagen. The mixture was incubated at room temperature for 5 to 10 minutes and added to the cells. Cells were fed with fresh glycine, hypoxanthine and thymidine-free (GHT-free) medium and twenty-four hours later, were trypsinized and selected in fresh GHT-free medium with 0 – 5 nM of methotrexate (MTX) in order to select for stable DHFR+ clones. Approximately 300 – 400 individual clones were selected in this first round of screening for measurement of protein expression levels. Clones from each vector which expressed the highest levels of antibody were then re-exposed to higher levels of methotrexate to affect a second round of gene amplification and selection. The screening process was repeated on all available clones, the highest of which were exposed to a third round of amplification. The methotrexate concentrations used during amplification using the pSV.ID-derived vector was 50 to 1000 nM in the 2nd round and 200 to 1000 nM in the 3rd round. These concentrations are typically required to achieve growth-limiting toxicity, which is required to achieve sufficient selective pressure for gene amplification. Concentrations required to reach this same degree of toxicity using the pSV.ID-derived vectors were remarkably lower.

The level of antibody expression was determined by seeding cells in 1 ml of serum-free F12:DMEM-based media supplemented with protein hydrolysate and amino acids in 24 well dishes at 3×10^6 cells/ml or in 100 ul of similar media in individual wells of a 96 well plate. Growth media was collected after 3-4 days and titers were assayed by an ELISA directed towards the intact IgG molecule. In experiments where cells were not seeded at equal cell densities, a fluorescent measure of viable cell number was performed on each well in order to normalize expression data. An Intact IgG ELISA was performed on microtiter plates which used a capture

antisera directed to framework Fab residues common in both antibodies. Media samples were added to the wells followed by washing and a horseradish peroxidase labeled second antibody directed towards common framework Fc residues was used for detection.

Table 2 presents expression level distributions of clones isolated during each round of screening of anti VEGF clones, which resulted from transfection with the plasmid containing only the DHFR sequence in the intron (pSV.ID.aVEGF), and 2C4 clones that were created using the Puro/DHFR fusion sequence in the same intron (pSV.IPD.2C4). The distribution of expression levels seen in the case of anti VEGF is typical of the performance of the vector containing only the murine DHFR gene in the intron (pSV.ID). All isolates identified in the first and second rounds of screening have relatively low expression levels. In the initial selection round, no clones with expression above 5 were isolated. At least three rounds of amplification are required to identify clones capable of specific productivity greater than 50. The 2C4 clones were screened after the first exposure to methotrexate (0-2.5 nM) and the most productive of these were exposed to a second round of amplification in 10-25 nM MTX. Cells surviving this amplification were pooled and exposed to 3rd round amplification prior to selection for further screening. In contrast to the pSV.ID vector, using the pSV.IPD vector, clones with an expression level of up to 25 were identified even in the first round of screening. Clones with an expression level greater than 25 represented 95% of the population after their third round of amplification and screening.

The data from Example 1 indicates that use of the Puro/DHFR fusion protein as the selectable marker allows for faster, more efficient isolation of highly productive CHO clones using significantly lower levels of methotrexate. The data suggests that exposure to low concentrations and stepwise increments in methotrexate allow for the efficient initial selection of highly expressing clones and subsequent gene amplification. Exposure to excessively high concentrations of methotrexate or large incremental increases in exposure often does not yield increases in gene expression since cells rapidly acquire methotrexate resistance through non-gene amplification mechanisms. Importantly, the data also shows that the Puro/DHFR fusion protein provides an unexpectedly impaired activity of the DHFR gene product or an enhanced sensitivity to methotrexate, which results in a highly stringent initial selection step, and allows efficient gene amplification at concentrations of methotrexate not frequently associated with the acquisition of drug resistance through alternative mechanisms. The ability to select cells which have incorporated the plasmid either in the presence of puromycin or methotrexate, prior to initiating exposure to

methotrexate also provides a means of transferring this efficient system to DHFR (positive) host cells.

For Example 1 the structure of the expressed antibody has been extensively characterized. The proteins generated from the pSV.IPD are indistinguishable from the antibody produced by the pSV.ID vector, with no apparent increase of free heavy or light chain expressed by the pool.

TABLE 2. PERCENTAGES OF pSV.IPD.2C4 CLONES ISOLATED AT VARIOUS EXPRESSION LEVELS AFTER MTX EXPOSURE¹

Expression Level ²	pSV.ID.aVEGF 1st Rd	pSV.IPD.2C4 1st Rd	pSV.ID.aVEGF 3rd Rd	pSV.IPD.2C4 3rd Rd
<1	71	16	0	0
1-5	29	67	0	0
5-10	0	14	2	3
10-25	0	3	15	4
25-50	0	0	35	21
50-100	0	0	46	61
100-150	0	0	2	3

¹MTX concentration for Control SD vector = 0-10 nM 1st round, 50 –1000 nM 2nd round, 200-1000 nM, 3rd round. SD- Puro/DHFR vector = 2.5 nM 1st round, 25 nM 2nd round, 100 nM 3rd round.

² Expression levels are in mg/ml or (mg/ml)/Fluorescent Unit

This example demonstrate the general applicability of the Puro/DHFR fusion sequence for selection of highly productive recombinant cell lines following minimal exposure to MTX.

EXAMPLE 2

Recombinant protein production using a pSV.I construct containing DHFR and a fusion gene other than Puro

Constructs can also be produced that contain a fusion sequence of an alternative selectable marker and DHFR within an intron region as described in Example 1. For instance

starting with the vector pSVID, the coding sequences for the neomycin resistance gene (Neo), hygromycin resistance gene (Hygro), glutamine synthase (GS), thymidine kinase (TK) or zeocin (Zeo) could be inserted in frame with the start site of the murine DHFR sequence contained within the intron. The stop codon of this inserted gene would then be removed using site directed mutagenesis according to example 1. Depending upon the phenotype of the host cell selected, cells incorporating the plasmid could then be selected using either GHT-free or MTX containing media as described in examples 1-3 or using an appropriate quantity of the alternative selective agent. Gene expression by the resulting clones could then be amplified in the presence of increased levels of methotrexate.

EXAMPLE 3

Direct Selection with plasmids SV.IPD.HP and CMV.IPD.HP after spinner transfection

DP12 CHO cells were grown in growth medium with 5% FBS (fetal bovine serum) and 1X GHT (glycine, hypoxanthine and thymidine). The process typically took about 4 days. On day 1, cells were seeded at 4×10^5 /ml in 400 ml growth medium in a 500 ml spinner vessel and grown for 2 days at 37 °C. On day 3, the exponentially grown cells were seeded at 1.5×10^6 cells/ml in a 250 ml spinner vessel containing 200 ml of growth medium plus 5% FBS and 1X GHT. The cells were grown for 1 to 2 hours at 37 °C before transfection. During that time, serum-free growth medium and 1X GHT was warmed to 37 °C. 400 µg plasmid construct DNA and 1 ml of Lipofectamine 2000[®] (Qiagen) were separately diluted into 25 ml of warm serum-free medium in 50 ml Falcon tubes. The solutions in the tubes were combined and incubated at room temperature for 30 minutes. The cells were then transfected with plasmid constructs pSV.IPD.HP and pCMV.IPD.HP, which constructs are illustrated in Figures 13 and 14, respectively. At the end of incubation, the cells were transfected by adding all 50 ml of the mixture of diluted plasmid construct and Lipofectamine 2000[®] to the 250 ml spinner vessel containing cells in serum-free medium, and the cells continued to grow at 37 °C for about 24 hours. On day 4, 250 ml of transfected cells were centrifuged at 1000 rpm for 5 minutes to collect the pellet. The transfection efficiency was monitored by transfecting cells with a GFP plasmid followed by FACS analysis 24 hours after transfection. The transfection efficiency with this protocol was typically approximately 55 to 70 % in CHO cells as shown in Figure 8.

After the transfection, cells were centrifuged to collect the pellet. The pellet was then resuspended in growth medium containing methotrexate (MTX) ranging from 10 to 100 nM for either SV40 or CMV based constructs. Approximately 100 clones survived the direct selection. Cell growth medium was changed every 3 to 4 days. At approximately 2 weeks after transfection, individual clones were picked and grown in 96-well plates in growth medium containing MTX. Heterologous polypeptide expression levels were evaluated by ELISA. Figures 10-1, 10-2, and 11 show the results from 25 nM and 50 nM MTX selection. Figure 9 shows heterologous polypeptide expression levels of clones from a traditional 10 nM MTX selection where the cells were not transfected in a spinner flask.

It took about 1 week for cells to grow confluent in a 96-well plate. When they were confluent, the growth medium was removed and commercially available enriched cell culture medium (which includes 1x GHT but no MTX) was added into each well. On the day after adding the commercially available enriched cell culture medium, the plate was incubated at 33 °C for 5-6 days before performing an ELISA assay to quantitate the amount of humanized monoclonal antibody produced by the cells. ELISA was typically performed with serial dilutions of the commercially available enriched cell culture medium. Results from a humanized monoclonal antibody production were shown in Figures 9, 10-1, 10-2 and 11.

The four clones producing the greatest amount over 100 µg/ml of intact IgG based on direct selection at 25 nM MTX using a CMV-based construct were scaled up from a 96-well plate to a 6-well plate and then to a 10 cm plate. Cells were seeded at 3×10^5 /ml in 200 ml volume in a 250 ml spinner vessel in serum-free growth medium with 2 µg/ml human insulin and 1X Trace Elements (TE). Cells were initially passaged at either two- or three-day intervals with medium exchange. Then they were passaged at either three- or four-day intervals for about 6 weeks before bioreactor evaluation. At each passage time, cell viability and count number were monitored. To determine the cell growth after serum-free adaptation, a spinner vessel growth experiment was performed. Cells were seeded at 3×10^5 cells/ml into 400 ml of growth medium with 2 µg/ml recombinant human insulin and 1X TE in a 500 ml spinner vessel on day 1. On each day, packed cell volume (PCV) was monitored until day 5. PCVs reached between 0.4 % to 0.6% by day 4. Two serum-free adapted clones from 25 nM MTX selection with CMV-based construct were evaluated in bioreactors. Two liter bioreactors with commercially available

enriched cell culture medium were run for a total of 14 days. The data from the titer evaluation is shown in Figure 12.

An ELISA assay of clones surviving the direct selection shows that the best clones coming out of the method described in this example produce as much product of interest as highly amplified clones from a traditional method. See Figure 16. Evaluations of 2 clones from the direct selection shows that those clones produce about 1g/L of a product of interest in a bioreactor process. Since those clones were generated from one step of a direct selection immediately after transfection, it only takes about 5 to 6 weeks to generate a stable cell line producing 1g/L of a product of interest in a bioreactor leading to significant timeline reduction, about 3 months, which is critical for efficiency of product development.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the examples presented herein, since the exemplified embodiments are intended as illustrations of certain aspects of the invention and any functionally equivalent embodiments are within the scope of this invention. The examples presented herein are not intended as limiting the scope of the claims to the specific illustrations. Indeed, various modifications of the invention, in addition to those shown and described herein and which fall within the scope of the appended claims, may become apparent to those skilled in the art from the foregoing description.

CLAIMS

What is claimed is:

1. A method of producing a host cell capable of producing a product of interest, comprising:

transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;

directly culturing the transfected host cells in a selective medium;

allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur; and

selecting a host cell clone that is capable of producing at least about 250mg/l of the product of interest.
2. A method of claim 1 wherein the selective medium contains at least about 25nM methotrexate.
3. A method of claim 1 wherein the selective medium contains at least about 50nM methotrexate.
4. A method of claim 1 wherein the host cell is a CHO cell.
5. A method of claim 1 wherein the product of interest is a protein selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin; or a fragment of said protein.
6. A method of claim 1 wherein said product of interest is a humanized antibody.
7. A host cell produced according to the method of claim 1.

8. A method of producing a product of interest, comprising culturing a host cell produced according to the method of claim 1 under conditions suitable to cause expression of the product of interest in an amount at least about 250mg/l.

9. A method of claim 1 wherein the DNA construct comprises, in order 5' to 3':

a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;

b) a transcriptional initiation site;

c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;

d) a product gene encoding a product of interest; and

e) a transcriptional termination site.

10. The method of claim 9 further comprising recovering the product of interest from the culture.

11. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a SV40 promoter.

12. A method of claim 9 wherein the transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene is driven by a CMV promoter.

13. A cell culture composition comprising a host cell according to claim 9 and at least about 250mg/l of the product of interest.

14. A method of producing a host cell capable of producing at least about 250mg/ml of a product of interest comprising transfecting a host cell with a DNA construct comprising in order from 5' to 3':

a) a transcriptional regulatory region capable of regulating transcription of both the selectable gene and the product gene;

b) a transcriptional initiation site;

c) a fused selectable gene sequence positioned within an intron defined by a 5' splice donor site comprising a splice donor sequence such that the efficiency of splicing messenger RNA having said splice donor sequence is between about 80% and 99% as determined by PCR, and a 3' splice acceptor site;

d) a product gene encoding a product of interest; and

e) a transcriptional termination site;

wherein the transfection is performed in suspension culture.

15. A method of claim 14, wherein the DNA construct is introduced into the host cells by lipofection.

16. A method of claim 14 wherein said transfection is performed in a spinner vessel.

17. The method of claim 14 wherein the suspension culture has cell density of at least about 5×10^5 /ml at the time of transfection.

18. The method of claim 14 wherein the suspension culture has a cell density of at least about 1.5×10^5 /ml at the time of transfection

19. A method of claim 15 wherein the product of interest is selected from the group consisting of an antibody, enzyme, hormone, lipoprotein, clotting factor, anti-clotting factor, cytokine, viral antigen, chimeric protein, transport protein, regulatory protein, homing receptor, and addressin and a fragment of any of said product of interest.

20. A method of rapidly selecting a host cell producing a product of interest, comprising:

transfecting a host cell culture with a DNA construct comprising a transcriptional regulatory region, a fused selectable gene sequence and a gene encoding a product of interest;

directly culturing the transfected host cells in a selective medium; and

allowing the host cells to grow in the selective medium for a sufficient time to allow amplification of gene encoding the product of interest to occur.

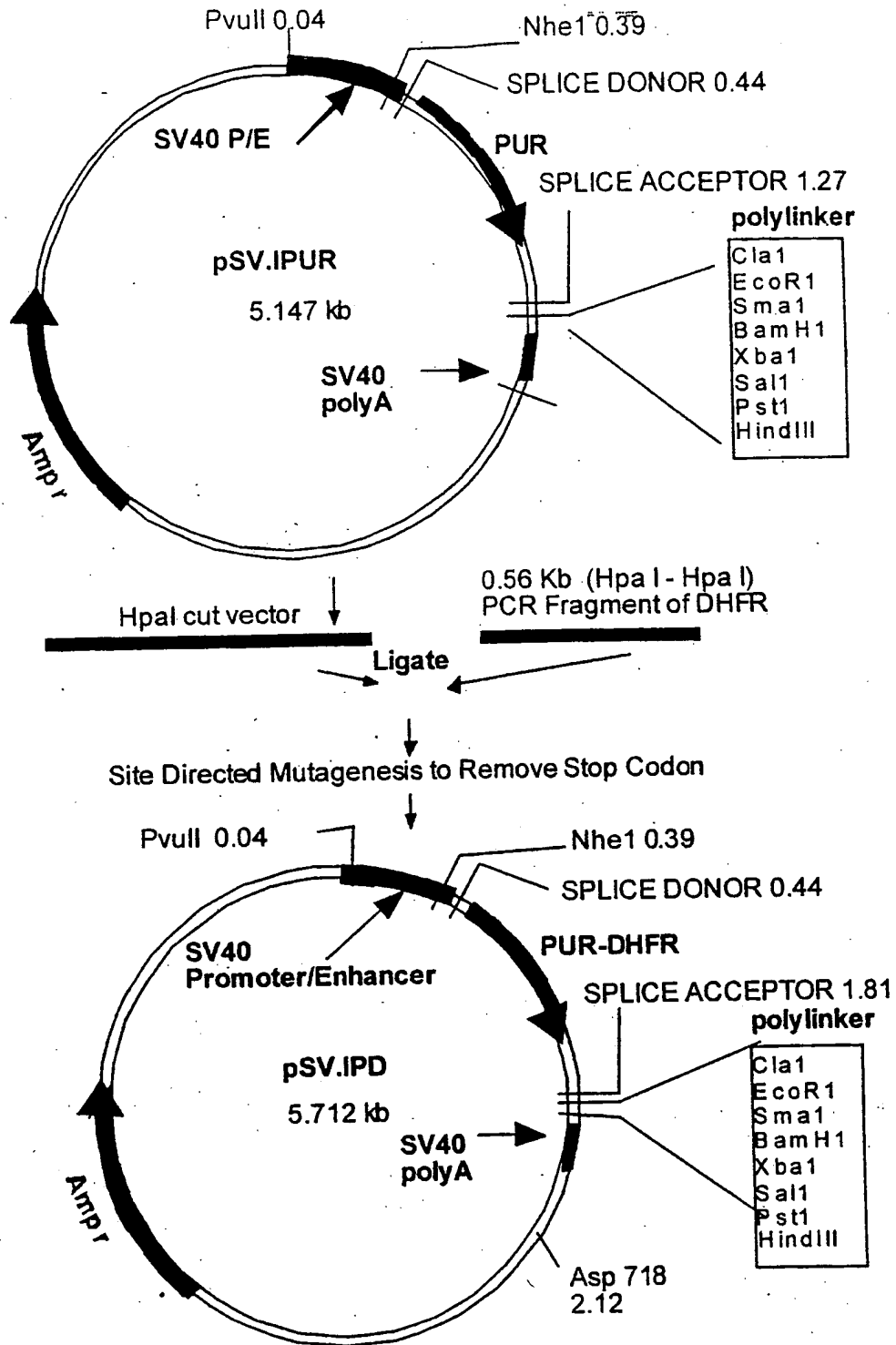


Figure 1. Construction of pSV.IPD Plasmid

Figure 2
psv.IPUR
length: 5147 (circular)

1 TTCGAGCTCG CCCGACATTC ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT CAGTTAGGGT GTGGAAAGTC CCGAAGCTCC CCAACACACAA
AAGCTCGAGC GGGCTGTAACT TAATAACTGA TCTCAGCTAG CTGTCGACAC CTTCACACA GTCAATCCCA CACCTTTCAG GGGTCCGAGG GGTTCCTCCT

101 GAAGTATGCA AAGCATGCAT CTCATTAGT CAGCACACCG GTGTGAAAG TCCCCAGGCT CCCGAGCAAG CAGAAGTATG CAAACCATGC ATCTCAATTA
CTTCATACCT TTCGTACGTA GAGTTAATCA GTGGTTGGTC CACACCTTTC AGGGTCCGA GGGTCTCTCC GTCTTCATAC GTTTCGTACG TAGAGTTAAT

201 GTCAGCAAC ATAGTCCCGC CCCTRACTCC GCGCATCCCG CCCTRACTC CGCCAGTTC CGCCATTCT CGGCCCATG GCTGALTAAT TTTTATTATT
CAGTCGTGG TATCAGGGCG GGGATTGAGG CCGGTAGGCG GGGGTTGAG CGGGTCAAG CGGGGTAC GACTGATTA AAAAAATAA

301 TATGCAGAG CCGAGCCCGC CTCGGCTCT GAGTATTCC AGAAGTAGTG AGGAGGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAAGCTA GCTTATCCCG
ATACGCTCC GGTCCGGCG GAGCCGGAGA CTCGATAAG TCTTCATCAC TCCTCCGAA ABACTCCGG ATCCGAAAC GTTTTCCGAT CGAATAGGCG

401 CCGGGAACGG TGCATTGGAA CGCGATTCC CGGTGCAAG AGTGCATAA GTACGGCTA TAGAGGACT AGTCACCAT GACCGAGTAC AAGCCACACG
GGCCCTTGC ACCTAACCTT GCGCTAAGG GGCACGGTTC TCACTGCATT CATGGCGAT ATCTCGTGA TCATGTA CTGGCTCATG TTTGGCTTCC

501 TCGGCTCGC CACCCCGCAC GACCTCCCGC GGGCCCTPAG CACCTCGCC CGCGGTTG CGACTACCC CGCCACGCGC CACACCTTCC ACCCAGTAC
ACCGGAGCG GTGGGGCTC CTCAGGGGG CCGGGATGC GTGGGCAAG GGTGATGG GCGTGCAGG GCGTGCAGG GCGTGCAGG GCGTGCAGG GCGTGCAGG

601 CCACATCGAG CCGGTACCG AGCTCAAGA ACTTCTCTC ACGGCGTGG GGTCTGACAT CGGCAAGGTG TGGTTCGCG ACAGGGGCG CAGTGCAGG
GGTGTAGCTC GCGCAGTGC TCACAGTCT TGAAGAGG TGGCGGAGC CCGAGTGA GCGTTCAC ACCCAGCGC TGCTCCCGC GCGTGCAGG

701 GTCAGACCA CCGCGGAGG CACTCGAGG GGGGGTGT TCGCCGAGT CCGCCCGCG ATGGCCGAT TGAGCGTTC CCGGTGCGC GCGCAGTAA
CAGCCCTGT GCGGCCCTC GCAGTTCCG CCGCGCAC AGCGCTCA CCGGGCGG TACCGCTCA ACTCCCAAG GCGCGACCG CCGTGCAGG

801 AGATGGAAG CTCCTCGCG CCGCACCGC CCAAGAGCC CCGTGTTC CTGSCCACC TCGCGCTCT CCGCGACCA CAGGCAAG GCTTGCAGG
TCTACCTCC GGAGGACCG GCGTGCAGG GGTTCCTCG GCGACCAAG GACGGTGGC AGCCGAGG CCGGCTGGT GTCCCGTCC CAGTGCAGG

901 CGCGTCTG TCCCGGAG TGGAGGCGG CAGCGCGCC GGGTGCAG CTTTCTGGA GACTCCCGC CCGCGAAC CCGCGAAC TCCCTTCTA CAGCGCTTC
GGGCGCAC GAGGGCTC ACCTCCCGC GCTCGCGG CCGCACGGC GAAAGACT CTGGAGGCG CCGCGAAC GTGGAGTTC AGGGAAAGT GCTCCCGAG

1001 GCGTTCACC TCACCCCGA CACTGAGTGC CCGAAGACC GCGGACTG GTGCATGACC CGCAAGCCCG GTGCTGAGT TAACGCTCC CAGTGCAGG
CGAAGTGC AGTGGGGCT GCAGTCAAG GCTTCTCG CCGTGCAG CACTGACC CCGTGCAG CAGGACTA ATTGACGAG GAGGAAATTC

1101 CTATGATTT TATAGACC ATGGACTTT TGCTGGCTTT AGATCCCTT GGTTCGTTA GACGCGACT ACAATTAATA CATAACCTTA TGTATCATAC
GATACATAA ATATTCTGG TACCCTGAA ACACCGAA TTAGGGAA CCGAAGAA CTGCGTGA TGTAAATAT GATTTGGAAT ACATACTATC

1201 ACATAGGATT TAGGTGACAC TATAGATAAC ATCCACTTTG CTTTCTCT CACAGGTGTC CACTCCCGG TCCCACTGCA CTTCTCTCT ATCCATTTAA
TGTATGCTAA ATCCACTGTG ATATCTATT TAGTGAAC GGAAGAGAG GTGTCCACAG GTGAGGTTCC AGGTTGAGCT GAGGCCAAGA TAUCTAATTT

1301 TTCCCGGGG ATCCTCTAGA GTCGACCTGC AGAAGCTTC ATGGCCCGCA TGGCCCACT TGTATTTCG ACTTATTAAT GCTTACAAAT AAGGCAATAG
AAGGGCCCC TAGGATCT CAGTGGAGC TCTTCBAGC TACCGCGGT ACCGGTTGA ACAATAACG TCGAATAATA CCAATGTTTA TTTCTTATT

Figure 2-1

1401 CATCAAAAT TTCACAAATA AAGCATTTTT TTCACATGCAAT TCTAGTGTGTG GTTGTGCCAA ACTCATCAAT GTATCTTATC ATGTCGTGAT CUATCGUGA
 GTAGTGTATA AAGTGTATTTT TTCGTAATAA AAGTGCAGTA AGATCAACAC CAAACAGGTT TGAGTAGTTA CATAGATATG TACAGACCTA UCTAGCCCTT

1501 TTAATTCGGC GCAGCACCAAT GGCCTGAAMT AACCTCTGAA AGAGGAACCT GGTAGTAGAC CTTCTGAGGC GGAAGAACC AGCTGTGAAA TGTGTGTATG
 AATTAAGCCG CGTCTGTGGTA CCGGACTTTA TTGGAGACTT TCTCCTTGAA CCAATCCATG GAAGACTCCG CCTTCTCTGG TCGACACCTT ACACACACTC

1601 TTAGGGTGTG GAAATGCCC AGGCTCCCCA GCAGGCAGAA GTATGCARAG CATCATCTC AATATGTCAG CAACCAGGTG TGAAGAATCC CAAGGCTCCC
 AATCCACAC CTTTCAGGGG TCCGAGGGT CCTCCCTCTT CATACCTTC GTACGTAGAG TTAATCAGTC GTTGGTCCAC ACCTTTCAGG GGTCCGAGGG

1701 CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACATA GTCCCGCCCT TAACTCCGCC CATCCCGCCC CTAACCTCCG CCAATTCGCG
 GTCGTCCGTC TTCATACGTT TCGTAGGTAG AGTTAATCAG TCGTTGGTAT CAGGGCGGGG ATTGAGGCGG GATTGAGCGG GGTCAAGGCG

1801 CCAATTCCTG CCCCATGGCT GACTAATTTT TTTTATTTAT GCAGAGCCCG AGGCCGCTC GGCTCTGAG CTATTCAGA ACTAGTGAAG AUCCTTTTTT
 GGTAAAGGC GGGGTACCGA CTGAATANA NAATAAATA CBTCTCCGC TCCGGGGAG CCGGAGACT GATPAGGTCT TCAATCACTCC TCCGAAAAA

1901 GGAGCCCTAG GCTTTTGCAA AAAGCTGTTA CCTCGAGCGG CCGCTTAATT AAGCCGCGCC ATTTAATCC TGCAGGTAAC AGCTTGGCAC TGGCGSTCGT
 CCTCCGGATC CGAARACGTT TTTCGACAAI GAGCTCGCC GCGGAATTA TTCCCGCGG TAAATTTAGG ACGTCCATTG TCGAACCGTG ACCGGCAGCA

2001 TTTCAACGT CGTAGCTGG AAAACCTGG CGTTACCCAA CTTAATCGCC TTGCAGACA TCCCCCTTC GCCAGCTGGC GTRATAGCGA AGAGCCCGC
 AARTGTGCA GCACGTACCC TTTTGGACC GCARTGGTT GAATFAGCGG AAGCTGTGT AGGGGGAAG CCGTGCAGCC CATTATCGCT TCTCCGGGCG

2101 ACCGATCGC CTTCCCAACA CTTGCGTAG CTGARTGGG AATGGGCTT TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACUUCATAC
 TGGTAGCGG GAAGGTTGT CAACGATCG GACTTACCGC TTACCGGGA CTACGCCA AAAGAGGAT GCTAGACAC GCCATAAAGI GTGGGTATG

2201 GTCARAGCA CCAATAGTAG CCCCCTGTAG CCGCGATTA AGCCGGCGG GTGTGTGTGT TAGCGGAGC GTACCGCTA CACTTGGCAG CUCCTTAGUG
 CAGTTTCTTT GGTATCATG CCGGACATC GCGCGCTAAT TCGCGCGCC CACACACCA ATCGCGCTG CACTGGCAT GTGAACGCTC GCGGATCGC

2301 CCGGCTCCTT TCGTTTCTT CCGTCCCTT CTCGCCAGT TCGCCAGCT GCGCGCTT TCCCGCTCAA GCTCTAAATC GGGGGCTCCC TTTAGGTTTC GATTTTATG
 GGGCGAGAA AGCGAAGAA GCGGCTGA GAGCGCTGA AGCGCCGAA AGGGCGAGT CGAGATTTAG CCCCCAGGG AATCCCAAG GCTAAATCAC

2401 CTTTACGCA CCTGACCC AAAAATCTG ATTTGGTGA TGGTTCACGT AGTGGCCAT CCGCCCTGATA GACGGTTTTT CGCCCTTGA CGTGTGATC
 GAAATGCCG TGAGCTGGG TTTTGTGAC TAAACCACT ACCAAGTGA TCACCCGGTA GCGGGACTAT CTGCCAAAA GCGGAAACT GCAACCTCAG

2501 CACGTTCTTT AATAGTGGAC TCTTGTCCA AACTGAAAC ACACCAACC CTAATCTGGG CTAATCTTTT GATTTATAAG GAATTTTGGC GATTTCTGGC
 GTGAGAAA TTATCACCTG AGACAAAGT TTGACCTTGT TGTAGTTGG GATGAGCCC GATGAGAAA CTAATATTC CCTAARACGG CTAAGGCCG

2601 TATTGTTAA AAAATGAGCT GATTTAACAA AAATTTAAG CGAATTTTA CAAATATTA AGTTTACAA TTTTATGGTG CACTCTCAGT ACAATCTCT
 ATACCAAT TTTTACTCGA CTAAATGTT TTTAATGCG GCTTAAATTT GTTTTATAAT TGAATACAC GTGAGAGTCA TGTATAGACA

2701 CTGATGCCC ATAGTTAAG CAACTCCGT ATCGCTACT GACTGGTCA TGGTCCGCC CGACACCCG CCAACACCCG CTGACCGCC CTGACCGGCT
 GACTACGGC TATCAATTC GTTGGGCA TAGCGTGA CTGACCACT ACCGAGCGG GGTGTGGC GACTGGCGG GACTGGCCG

2801 TGTCTGTC CCGATCCGCT TTACAGCAA GCTGTACCG TCTCCGGGAG CTGCAATGTT CAGAGTTTTT CACCGTCAIC ACCGAAACG GUAAGUAGT
 ACAGACGAG CCGGTAGCGC AATGCTGTT CGACACTGC AGAGGCCCTC GACTGACACA GTCTCCAAA GTGGCAGTAG TGGTTTTGCG CCGTCCCTA

2901 ATTCTGTAG ACGAAAGGC CTCGTGATAC GCTATTTTT ATAGTTAAT GTCAIGATAA TAATGTTTT TTAGAGTCA UGTGUKACTT TTCTATKANA
 TAAGRACTTC TGCYTTCCG GAGCACTATG CCGATPAAA TATCCAAITA CAGTACTAT ATTACCAAG AATCTCAGT CCACTUJAA AUGUULYTTT

3001 TGTGCGGGA ACCCCTATTT GTTTATTTTT CTAAATACAT TCAATATGAT ATCCGCTCAT GAGACAATAA CCCTGATMAA TGGTTCARATA ATATTCANAA

Figure 2-2

ACRCCGCCCT TGGGATATA CAATRAAAA GATTTATGTA AGTTTATACA TAGCGGAGTA CTTCTGTATT GGGACTATTT ACGAAGTTAT TATAACTTTTT
3101 AGGAAGATA TGAGTATTCA AGATTCCCGT GTCCGCCHTA TTCCCTTTTT TGGGSCATTT TCCCTTCCG TTTTGTCTCA CCCASAAAACG CTUUTUAAAAG
TCCTTCTCAT ACTCATRAGT TGTAAAGGCA CAGCGGGAAT AAGGNNAAA ACGCCSTRAA AGGRRAGAC AAAACRAGT GGTCTTTCG GACCACCTTC
3201 TAAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGGTTA CATCGAATCG GATCTCRAA CCGGTAAGT CCTTCAGAGT TTTTGGCCCGG AAGAALGTTT
ATTTTCTAGC ACTTCTAGTC AACCCACGTC CTCACCCAA CTACCCCAAT GTACCTTGAC CTAGAGTGT GCCCATCTA GGAATCTCA AAAGCCGGUC TTCCTTCCAAA
3301 TCCATATGAG AGCACTTTTA AAGTCTTGCT ATGTGGCGG GTATTATCCC GTGATGACGC CCGCAAGAG CAATCGGTC CCGCATPACA CTAATCTCAG
AGGTACTAC TCGTGAAT TCAAGACGA TACRCCGGC CATRATAGG CACTACTGGG GCCCTTCTC GTTCAGCCAG CCGCTTACT GATPAGACTC
3401 ARTGACTTGG TTGAGTACTC ACCAGTCACA GAAAGACTC TTACGGATGG CATGACAGTA AGGAATAT GCACTGCTGC CATAACTAG AGTATATAA
TTACTGAAAC AACTCATGAG TGGTCAATGT CTTTTCGTAG AATGCCACC GACTCTCAT TCTCTPATA CACTCAGAG GTATTGCTAC TCACATATTT
3501 CTGGCCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGACTAACG GCTTTTTTG ACAACATGG GATCATGTA ACTTGGCTTG ATCTTCAKKA
GACGCCGTT GAATGAGAC TGTGCTAGC CTCCTGGCTT CCTCGATTG CGMAAAAAG TGTCTTACC CTPACTACAT TGAGCCGAC TACCAACCTT
3601 ACCGAGCTG AATGAGCCA TACCAACGA CGAGCTGAC ACCAGATGC CAGCAAT GCGCAACG TTGGCCAAAC TATTAACCTG CCAACTATTT
TGGCTCGAC TTACTTGGT ATGGTTTGT GCTCCACTG TGGTCTACC CTCTCTTA CCGTGTTCG AACGGTTG APARTTGACC SCTTGTATTA
3701 ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAG CGGATAAGT TGCAGACCA CTTCTGCGT CCGCTGCTCC GCGCCGAGG CCGACCGACC AAATRALAC
TGGATCGAA GGGCCGTTG TAATTATCG ACCTACTCC GCTATTTCA ACGTCTGTG SAAGACGGA CCGCCGAGG CCGCCGAGG CCGACCGACC AAATRALAC
3801 ATAAATCTGG AGCCGCTGAG CBTGGCTTC GCGTATCAT TGCAGACTG GGGCCAGATG STAAGCCCTC CCGTATGTA GTTATCTACA LUALUULAG
TATTPAGACC TCGGCCACTC GCACCCAGAG CGCCATAGTA ACSTGAGAC CCGGCTTAC CATTCGGGAG GGCATAGCAT CAATAGATCT CCTTCCUUTG
3901 TACGCRAT ATGATGAC GAAATAGACA GATCGCTGAG ATAGTGCCT CACTGATTA CCAITGTTAA CTCTCAGACC AAGTITACTC ATATATATTT
AGTCCGTTGA TACCTACTG CTTTATCTGT CTAGGACTC TATCCRGGG CTCACATAAT COTRACCATT GACACTCTG TTTCAAAATGAG TATATATUAA
4001 TAGATGAT TAAACTTCA TTTTAAATTT AAAAGATCT AGGTGAGAT CTTTTTCTA GCAAAATPCA TTAGAGTACT GCTTTTAGG AATTCACATC TTTTCTGTTA
ATCTAACATA ATTTTGAAT AAAATTTAAA TTTTCTTGA TCCACTTCTA GCAAAATPCA TTAGAGTACT GCTTTTAGG AATTCACATC TTTTCTGTTA
4101 ACTGACGCTC AGACCCCGTA GAAAGATCA AAGGATCTC TTGAGATCT TTTTTCTGC CCGTAATCTG CTGCTTGCAG ACAAATAAC CACCTTATCT
TGACTCGCAG TCTGGGCTT CTTTTCTAGT TTCTTGAAG AACTCTAGGA AAAAAAGAG CCGATTAGC GAGCAACGTT TGTTTTTTTG STRAUCATUJ
4201 AGCGGTGTT TGTTCGCGG ATCAGAGCT ACCAATCTT TTTCCGAGG TACTGGCTT CAGCAGAGC CAGATACCAA ATACTGTCTT TCTAGTUTAG
TCGCCCAA ACARACGGCC TAGTCTCGA TGGTTGAGA AAGGCTTCC ATTCACGAA STCCTCTCG GCTATGCTT TATGACAGCA AGATCACATC
4301 CCGTAGITAG GCCACACTT CAGACTCT GTAGACCGC CTACATACT CCGTCTGTA ATCCGTATC CAGTGGCTG TCGCAGTGC GATAATCTT
GGCATCATC CCGTGTGAA GTTCTTGA CATCTGCGG GATGATGA GCGACAGCAT TAGGACATG CTCACCGAG AGGCTCACG CTATTCAGCA
4401 GCTTACCGG GTTGGACTA AAGGATAGT TACCGATPA GCGCGAGCG TCGGGGTGA CCGCCCAAG CAGTGTGTC CCGCTTGC AKAJANAAJ
CAGAAAGCC CAACCTGAT TCTGTATCA ATGCTTAT ATGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT
4501 CTACACGAA CTGAGATACC CATTGAGGA GCAITGAGGA AGGCCACGC TTCCGAGAG GAGAAAGCG CAGAGATAC CCGTAAAGCG CAGATKHAJ
GATGTGGCTT GACTCTATGG ATGTGCACT CTTAATCTT CCGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT CCGCTTAT
4601 ACAGGAGC GCACGAGGA CTTCCAGG GAAACCCCT GATCTTPTA TACTCTGTC GGTTCGCT ACCCTGACT TGAUGTUGA TTTTPTTAT
TGTCTCTCG CBTCTCCCT CCAAGTCC CTTTGGGA CCAATAGAT ATCAGAGAG CCAAAAGCG TGAAGACTG ACTTCCAGT AAAAACTA
4701 CCTGCTCAG GGGCGGAGC CTATGAAA AGCCAGAA CCGGCTTCTT TACGGTTC TGGCTTCTG TGGCTTCTG TGGCTTCTG TGGCTTCTG
CGAGCATCC CCGGCTCG GATACCTTT TCGCTGTT GCGCCGAA ATGCGAAG ACCGAAAAC GACCCGAAA CUAUUNUJ

Figure 2-3

4801 GTTATCCCTT GATTCGTGG ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCCGCTCG CCGCAGCCGA ACGACCGAGC GCAGCGAGTC AATTGAGCCGAA
CAATAGGGGA CTAAGACACC TATTGGCATA ATGGCGGAAA CTCACCTCGAC TATGGGGAGC GGGCTCGGTG TGCTGGCTCG CFTCGCTCAG TCACTCGCTC
4901 GAAGCGGAAG AGCGCCCAAT ACGCAAACCG CCTCTCCCG GCGTTGGCC GATTCATTAA TCCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGCA
CTTCGCCCTT TCGCGGGTTA TGCCTTTGGC GGAGAGGGGC GCGCAACCGG CTAAGTAATT AGTTCGACCG TGCTGTCCAA AGGGCTGACC TTTCGCCCTT
5001 GTGAGCGCAA CGCAATTAAT GTGAGTTACC TCACTCAATTA GGCACCCGAG GCCTTACT TTAIGCTTCC GGCTCGTATG TTGTGTGAAA TTCTGAGCCG
CACTCGCGTT GCGTAAATTA CACTCAATGG AGTGAGTAAT CCGTGGGGTC CGAATGTGA AATACGAAGG CCGAGCATAC AACACACCTT AACACTCCGC
TATTGTTAAA GTGTGCTCCTT TGTCGATACT GGTACTAATG CTTAATT

>length: 5147

Figure 2-4

Figure 3
PSV.ID
length: 5171 (circular)

1 TTCAGACTCG CCCACAAATTG ATTAATGACT AGAGTGAATC GACAGCTGTG GAATGTGTGT CAGTTAGGTT GTGGAAAGTC CCCAGGCTCC CCACAGAGCCA
 AAGCTCGAGC GGGCTGTAAAC TAATAAAGTGA TCTCAGCTAG CTGTGACACAC GTCAATCCCA CACCTTTCAG GGGTCCGAGG GUTTCGTCCT

101 GAAGTATGCA AAGCATGCAT CTCAAATAGT CAGCACCAG GTGTGGAAAG TCCCAGGCT CCCAGCAGG CAGAAGCTATG CAAAGCAATCC ATCTCTATATTA
 CTTCAATAGT TTGCTAGCTA GAGTTAATCA GTCTGTGTGC CACACCTTTC AGGGTCCGA GGGTCTGTCC GTCTTCATAC GTTTCCTACG TAGACATTAAT

201 GTCACCAACC ATAGTCCCGC CCCTAACTCC GCCATCCCG CCCCTAACTC GCCCAGTTC CCCCATTCT CCCCCTCATG CCTGACTAAT TTTTATTATTA
 CAGTCGTGG TATCAGGGCG GGGATTGAG GGGATTGAG CCGGTCAAG CCGGTCAAG GCGGGGTAC CGACTGATTA AAAAAATAA

301 TATGCAAGG CCGAGGCGC CTCGGCTCT GAGCTATTCC AGAAGTAGTG AGGAGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAAGTA GCTTATCTGG
 ATAGCTCTCC GGTCCCGCG GAGCCGGAGA CTCGTAAGG TCTTCATCAC TCCTCCGAAA AAACCTCCGG ATCCGAAAAC GTTTTTCGAT CGAATAGGCC

401 CCGGACAGG TGCATTGGA CCGGATCC CCGTCCCAAG AGTACGTA GTCGCCCTA TAGAGTCTAT AGGCCACCC CTTGGCTCTA CAGAGATATA
 GGGCTTCC ACSTAACTT GCGCTTAGG GGCACGGTTC TCACTGCAT CATGGCGGAT ATCTCAGATA TCCGGTGGG SAACCGAGAT CTCTCTATATTA

501 AGCTTAGT TTTATCCCG GTGCCATCAT GGTTCGRCCA TTGAATGCA TCGTCCCGT GTCCCAAAAT ATGGGATG GCAAGAAGG ALACTATCTC
 TCGGATCTTA AATPAGGGC CACGGTAGTA CCAAGCTGTT AACTGAGGT AGCAGGGCA CAGGTTTTTA TACCCTTAAC CGTTCTTCC TCTGGGATGCG

601 TGCCCTCCG TCAGGACGC GTTCAAGTAC TTCCARCAA TGACCACAA CTTCTCAGTG GAAGTAAAC AGAATCTGTT GATTAAGGTT ALJAAAACCT
 ACGGGAGCG AGTCCCTTGC CAAGTTCAIG AAGTTTCTT ACTGGTGTG GAGAACTCAC CTTCCATTG TCTTAGACCA CTAATACCCA TCTTTTGGCA

701 GGTTCCTCAT TCCTGAGNAG AATCGACCTT TARAGACAG AATTAATATA GTTCTCAGTA GAGAACTCAA AGAACACCA CGAGGAGCTC AATTTCTTTC
 CCAAGAGGTA AGGACTCTTC TTAGCTGGAA ATTTCTCTGC TTAATATATA CAGAGTCAAT CTCTTGAGTT TCTTGGTGT GCTCTCCGAG TAAAAGAACU

801 CAAAAGTTG GATGATGCT TARACTTAT TGRACACCG GAATGGCAA GTRAGTAGA CATGGTTGG ATAGTCGGAG CCAGTCTCTG TPAALAGCAA
 GTTTTCAAAC CTACTAGGA ATTCTGAATA ACTTGTGGC CTTAACCTT CATTTCATCT GTACCAACC TATCAGCCTC CGTCAAGACA AATGGTCTCT

901 GCCATGAATC AACCAGCCA CTTAGACTC TTTGTGCAA GGAATTTGAA AGTACACGT TTTTCCAGA AATGATTTG GCGAATATA
 CGSTACTTAG TTGTTCCGGT GGAATCTGAG AACACCTGT CTTAGACTT CTTAARACTT TCACTGTGCA AAAAGGCTT TTAACATAAC CUCCTTATAT

1001 AACCTCTCCC AGAATACCCA GCGTCTCT CTGAGTCCA GGAGRAAAA GGCATCAAGT ATAAGTTTGA AGTCTACGAG AAGAAGACT AACAAGUAGA
 TTGGRAGGG TCTTATGGT CCGCAGGAGA GACTCCAGT CTTCTTTTT CCGTAGTTCA TATTCAAAT TCAGATGCTC TTTCTCTGA TTTCTCTTCT

1101 TGCTTTCAAG TTCTCTGCT CCCTCTAAA GCTATGANT TTTATAGAC CATGGACTT TTGCTGGCTT TAGACCCCT TAGCTTCTCT AGAALGCGAK
 ACGAAGTTC AAGAGACGAG GGGAGGATTT CGATAGTAA AATATTTG GTRCCTGAA AACGACCGAA ATCTGGGGA ACCGAGCAA TCTTTGCGLU

1201 TACAAATTAAT ACATACCTT ATGTATCATA CACATGAT TAGTGACAC TATAGATAA CATCCACTT GCCTTCTCT GCACACTGCT CATTTATAT
 ATGTTAATTA TGTATTGGAA TACATGAT GTGATCTTAA ATCCACTGTG ATATCTTAT GTAGTGAAA CCGAAAAGAGA GGTGULLACA CTGAGCTTCA

1301 CAACAGCACC TCGGTTCTAT CGATTGAAT CCCCGGGAT CCTTAGAGT GGCCTGAC AAGCTTGGCC CCAATGCCC AACTTCTTTA TTTAGACTTA
 GTTGACGTGG AGCAAGATA GCTAACTTAA GGGGCCCTA GGAGATCTCA GCTGRCCTC TTCGAACCG GGTACCUGG TTGAACAAAT AACTTCTAAAT

1401 TAATGTTTAC AAATAAAGCA ATAGCTCAC AATATTACA AATAAAGAT TTTTCTACT GCATCTAGT TGTGGTTGT CAAAALJLAT CANNVATATCT

Figure 3-1

ATTACCAATG TTTATTTTCGT TATCGTAGTG TTTAARAGTGT TTAFTTGGTA AAAAAGTGA CGTAAGATCA ACACCAAAACA GGTTCAGTGA GTTACATAGA
1501 TATCRTGTCT GGATCGATCG GGAATPRAAT CGGGGACGCA CAAFGGCGTG AAATAACCTC TGAAGAGGA ACTTGGTTAG GTACCTTCTG AGUCUGAAAU
ATAGTACAGA CCTAGCTAGC CTTAAATTA GCGCGTCTGT GGTACCGGAC TTTATTGGAG ACTTCTCTCT TGAACCAATC CATGGAGAC TCCGCTTTC
^sv40 origin
1601 AACCCAGCTGT GGAATGTGTG TCAAGTAGGG TGTGGAAGT CCCCAGGCTC CCCAGCAGG AGAAGTAGTC AAAGCATGCA TCTCAATPAG TCACGCAALCA
TTGGTCGACA CCTTACACAC AGTCAATCCC ACACCTTCA GGGTCCGAG GGTTCATRAC TTTCTCATRAC TTTCTCATRAC TTTCTCATRAC TTTCTCATRAC
1701 GGTGTGGAJA GTCCCAGGC TCCCAGCAG GAGAAGTAT GCNAGCATG CATCTCAAT AGTCAGCAAC CATAGTCCCG CCCCTAACCTC CCCCCTATCC
CCACACCTTT CAGGGTCCG AGGGGTCTG CGTCTTCAFA CGTTCTGTAC GTAGAGTAA TCAGTCTGTG GTATCAGGCG GGGATTTGAG GCGGGTAGUG
1801 GCCCTAACT CGCCCAATC CGCCCAATC TCCGCCCAT GGCCTGACTAA TTTTTTTTAT TTATGTCAG GCCAGGGCG CCI'CGGCCCTC TGAGCAATTC
CGGGATTA GCGGGTCAA GCGGGTAA GCGGGTAA GCGGGGTA CCGACTGAT AAAAAAAATA AATACGTCTC CCGCTCCGCG GGACCCGAG ACTCGATAAG
1901 CAGAACTAGT GAGGAGCTT TTTTGGAGC CTAGGCTTTT GCNARAGCT GTTACCTCGA CCGCCGCGCTT AATTAAGGCG GCCTAATTA ATCTGCAUG
GTCTTCAUCA CTCCTCCGAA AAAACCTCCG GATCCGAAA CTTTTTTCGA CAATGGAGCT CCGCGGGGAA TTAATTCGCG CCGCTAAATT TAGGACATTC
^start pUC118
^linearization linker inserted into HpaI site
2001 TAACAGCTTG GCATGCGCG TCGTTTTACA AGTCTGTGAC TGGGAARAC CTGCGCTTAC CCAACTTAAT CGCCTGGAG CACATCCCGC CTTCGCGCAUC
ATTFPCGAC CGTGACCGCG ACGAATATG TCGAGCACTG ACCCTTTTGG GACCGAATG GGTGAAATTA GCGGAACGTC GTGTAGGGGG GAAGCGGTGUG
2101 TGGCGTAATA GCGAAGAGC CCGCACCGAT CGCCCTTCCC AACAGTTGCG TAGCCTGAAT GCGCAATGCG GCCTGATGCG GTATTTTCTC CTTCACCAATC
ACCGCATPAT CGCTTCTCG GCGTGGCTA GCGGGAGGG TTGTCAACGC ATCGACTTA CCGCTRACCG CCGACTACCG CATAAAGAG GAATGCGTAG
2201 TGTGGGTAT TTCACACCG ATACCTCAA GCNACATAG TAGCGCCCTC TAGCGCGCG APTAAGCGCG CCGGGTGTGG TGGTTACGCG CAKCTMGAATC
ACACCCATA AAGTGTGGG TATGCAGTTT CPTTGTATC ATCGCGGGA CPTCGCGCG TAAITCGCG CCGCCACACC ACCAATCGCG GTUGACATUG
2301 CCTACACTG CCAGGCCCT AGCGCCGCT CTTTTCCCTC TCTTCCCTC AGTGTGCGG GAAAGAGCG AAGAGCGCG GCAAGAGGCG AGTTCGAGAT TTAGCCCLUL
CGATGTGAAC GGTCCGGGA TCGCGGGGA GAAAGCGAA AGAAGGGAAG GAAAGAGCG TCGAAGCGCG CCAAGTGGG CCATCGGCTT GATAGALUJ
2401 TCCCTTTAG GTCCGATTT AGTGTCTTAC GGCACCTCGA CCCCCAAAA CPTGATTTG GTGATGTTG ACCTAGTGG CCATCGGCTT GATAGALUJ
AGGAAATCC CAAGCTRAA TCACGAAATG CCGTGGAGCT GGGTTTTTTT GAACTAAACC CACTACAAG TGCATCACCC GGTAGCGGGA CTATCTGCA
2501 TTTTCGCCCT TTGACGTTG AGTCCAGCTT CTTTAAATG GACTCTTGT TCCAAACTG AACACACTC AACCCCTATCT CCGGCTATTC TTTTUAATTTA
AARAGCGGA AACTGCAAC TCAGTGCAA GAATATCA CCTGAGACA AGTGTGACC TTGTTGTGAG TTGGGTAGA CCCCATAAG AAAAACAAT
2601 TAAGGGATTT TGCCGATTC GGCCTATTG TTAARATG AGTGAATG AAAAAATTT AACCGAATTT TTAACAAAAT ATTAACGCTTT ACAATTTTAT
ATTCCCTAAA ACGGCTAAG CCGGATAAC AATTTTAC TCGCTRAAT TGTTTTTTAA TTGCGCTTAA AATTTGTTTA TAAITGAAA TGTAAAAATA
2701 GGTGACTCT CAGTACATC TGCTGTATG CCGCATAGT AAGCCACTC CGCTATCGCT ACCTGACTG GTCATCGCTC CCGCCCGACA CCCCCTAAC
CCAGTGAGA GTCATGTTAG ACGAGACTAC GCGTATCA TCCGTTGAG GCGATAGCGA TCGACTGACC CAGTACCGAC GCGGGGCTGT GCGCGGCTGT
2801 CCGCTGACG CGCCTGACG GCTTGTCTG CTCGCGGAT CCGCTTACG ACAAGCTGT ACCGTCTCG GAGCTGCAAT GTGTACAGAG TTTTACATJH
GGCGACTGC GCGGACTGC CCGAACAGAC GAGGCGCTA GCGGAATGTC TGTTCGACAC TGGCAGAGC CCTCGACTC CACAGTCTC AAAAGTGCBA
2901 CATCACGAA ACGCGGAGG CAGTATTTT GAAGCGARA GGGCTCTGT ATAGCCCTAT TTTTATAGGT TAAATGATG ATAAATAGG TTTTCTTAAJ
GTAGTGGCTT TCGCGGCTC GTCATAAGAA CTTCTGCTTT CCGGAGCAC TATCGGGATA AAAATATCCA ATTACAGTAC TAAATTTACC AAAGAATTC
3001 GTCAGGTGC ACTTTTCGG GAATGTGCG CGAACCCCT ATTTGTTTTT TTTTCTAAT ACATCAAT ATGTATCGC TCATGAGACA ATAAALCTGA
CAGTCCACCG TGAAGGCC CTTTACAGC GCCTTGGGA TAAACAATA AAAGATTTA TGTAACTTTA TACATAGGG ASTACTCTGT TATTTGCACT
3101 TAAATGCTC AATAATATTG AAAAGGAAG AGTATGATG TCAACATTT CCGTGTGCG CTTATTCCT TTTTGTGCG APTTTTCTT CTTTCTTTT
ATTACCGAG TTATTAAC TTTTCTTC TCATACTGT AAGTTGTAAA GGCACAGCG GAATAAGGGA AAAACGCG TAAAACGCAA GGCACAAA

Figure 3-2

3201 CTCACCCAGA AACGCTGGTG AABGTABABAG ATGCTGABAGA TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGGGGTA AATYTCYHGA
 GAGTGGGTCT TTGGACCAC TTTCATTTTC TACGACTTCT AGTCARCCCA CBTGTCCACC CAATPAGCT TGACCTAGAG TTGTCCCAT TCTAGGAACT

3301 GAGTTTTCCG CCGNAGAAC GTTTCCAAAT GATGAGCACT TTTAAAGTTC TGCTATGTTG CCGGTAATTA TCCCGTGTAG ACCLUGGCA AGAGCAACTC
 CTCAAAAGCG GGGTCTCTTG CANAAGTTA CTACTCGTGA AANITTCRAG ACGATACACC GCGCCATAAT AGGGCACTAC TCGGGCCCGT TCTCTGTTGAG

3401 GGTCCGCGCA TACACTATTC TCAGATGAC TTGGTTGAGT ACTCACCACT CACGARAAG CAVITPACGG ATGGCATGAC ACTAAGAGAA TPTATCCASTU
 CCAGCGGCT ATGTGATAAG AGTCTTACTG AACCAACTCA TGAGTGTCA GTGTCTTTTC GTAGATGCC TACCGTACTG TCATTCTCTT ATACGCTAC

3501 CTGCCATAAC CATGAGTGTAT AACACTGGG CCAACTTACT TCTGACRAG ATCGGAGGAC CGAAGAGGT AACCGGTTTT TTGSCACAACA TGGGGATCA
 GACGGTATG GTACTACTA TTGTGACGCC GGTGAAATGA AGACTGTCC TAGCTCTCTG GCTTCCGAAA AACGTGTGTG ACCCCGTAGT

3601 TGTAACTCGC CTTGATCGTT GGGAACCGGA CTTGAATGAA GCCATACAA ACGACRGG TGACACCAG ATGCCAGCAG CAATGCAAC AUCYTTCTUL
 ACATGAGCG GMACTAGCAA CCTTGGGCT CGACTTACTT CGTATGTTT TGTCTGTCG ACTGTGTGTC TACGGTCTG GTTACCCTTG TTGCAACGCG

3701 ABACTATTA CTGGGCACT ACTTACTCTA GCTTCCCGC AACATTAAT AGACTGGAG GAGGGGATA MAGTTCAGG ACCACTYCTG CUCYTCGGTCC
 TTTGATAAAT GACCGCTTGA TGAATGAGT CBRAGGGCG TTGTTAATTA TCTGACCTAC CTCCGCTAT TTCAACGTCC TGGTGAAGAC GCGAGCCCGG

3801 TTCGGGCTG CTGGTTTAT GCTGATTAAT CTGAGCCGG TGAGGGTGG TCTCCGGTA TCAITGACG ACTGGGGCCA GATGTPAAG CTTCCCGTAT
 AAGCCGACC GACCAATTA CGACTATTTA GACCTCGCC ACTCGACCC AGAGGCCAT AGTACGTG TGACCCCGGT CTACCATTCG GGAGGCGATA

3901 CGTAGTTATC TACACGAGG GGAGTCAGG AACATGAGT GAACGAATA GACAGATCG TGAGATAGT GCTCAGCTGA TTAAGCATTG GTAACCTCA
 GCATCAATAG ATGTGCTGCC CCTCAGTCC TTGATACCTA CTTGCTTAT CTGTCTAGG ACTCTATCCA CCGGTGACT AATTCTGTAAC CAITGACAGT

4001 GACCAAGTTT ACTCATATAT ACTTAGATT GAITTAAAC TTCTATTTTA ATTTAAAAG ATCTAGTGA AGATCCTTTT TGTAAVCTC ATUACUAAA
 CTGTTTCAA TGAATATA TGAATCTAA CTBAAATTTG ANGTAAAT TAATTTTCC TAGATCCACT TCTAGAAA ACTATTAGAG TACTGGTTTT

4101 TCCCTTAAAG TGAGTTTTG TTCCACTGAG COTCRAGCC CPTAGMAAG ATCAAAGAT CTTCTTGAGA TCCTTTTTTT CTGCGGTAA TCTGCTGCTT
 ABBGATTTG ACTCAAAGC AAGTGACTC GCAGTCTGG GCATCTTTTC TAGTTTCTTA GAAGAACTCT AGGAAAAAA GACGGGCAT AGACGACGAA

4201 GCAAAACAAA AAACCACCG TACCAGGGT GTTTTGTTG CCGATCAAG AGTACCAC TCTTTTTCCG AAGTACTG GCTTACGAG AGGCGAGATA
 CGTTTTGTTT TTTGGTGGG ATGTGCGCA CAAAACAAA GBCCTAGTTC TCGATGGTTG ARAAAPAGG TTCCATTGAC CGAAGTCTC TCGCTCTAT

4301 CCBARACTG TCCTTCTAGT GTAGCCGTAG TTAGCCACC ACTCAAGAA CTCGTGAGCA CCGCTACAT GCGCATCTA TGGAGCGAGA GATTAGGAC AATGCTCACC
 GTTTTATGAC AGAAGATCA CRTGGGATC AATCCGGTGG TGAAGTCTT GAGCATCCT GCGCATCTA TGGAGCGAGA GATTAGGAC AATGCTCACC

4401 CTGCTGCCAG TGGGATAAG TCFHTGCTTA CCGGTTGGA CTCAGAGGA TAGTTACCGG ATAAAGCCCA CCGTCCGGC TGAACGGGG GTTCTGTGCAC
 GACGAGGTC ACCGCTATTC AGCACAGAT GGCACACCT GAGTCTGCT ATCAATGSCC TATTCGCGT CCGCACCCG ACTTSCCCG CAAGCACGTC

4501 ACAGCCOAG TTGGAGGGA CGACTACAC CGAATGAGA TACCTACAGC GTGAGCATTG AGAAGGCC CCGTTCGCG AAGGAGAAA GCGGACAGG
 TGTCCGGTCC AACCTGCTT GCTGGATG GCTTGACTCT ATGGATGCG CACTGCTAAC TCTTCCGGG TCGAAGGGC TTCCCTCTTT CCGCTGTCC

4601 TATCCGATA GCGGAGGGT CCGAACAGGA GAGGACTCC AGGGGAAAC GCTGTGATC TTTATAGTCC TGTGCGGTT TGCACCTCT
 ATAGGCCAT CCGGTCCTA GCCTTGTCT CTGCGTCT CCCTCGAAG TCCCGCTTTG CCGACCATAG AATATCAG ACAGCCAAA GCGGTGAGA

4701 GACTGAGCG TCGATTTTT TGATGCTCT CAGGGGGCG GACCTATG AAAACCCCA GAAACCGGC CTTTTACGG TTCTGSGCT TTTCTGCTC
 CTGAACTCG AGCTAAAAC ACTACGAGCA GTCCCCCGC CTCGGATACC TTTTTGGGT GFTTGGCCG GAAAAATGCC AAGGACGGA AAAGUCLUG

4801 TTTTGTCTAC ATGTTCTTTC CTGCTTATC CCCTGATCT GTGATPACC GTATTACCG CTTTGTGTA -CTGTATCCG CTCGCGGAG CTAJAAVJAC
 AAAACGAGT TACAAGAAAG GACGCAATG GGCATATGG CAATATGGG GAAACTCACT CGACTATGG GAGCGGCTC GAACTTCTA TGG

4901 GAGCGCAGG ATCTAGTAG CAGGAAAGG GAAGAGGCG CAATACGAA ACCGCTCTC CCGCGGCTT GCGGATTC TTAATCCAGC TGUACUJAV

Figure 3-3

```
CTCCGGCTCGC TCAGTCACTC GCTCCTTCGC CTTCTCGCGG GTTAIGCGTT TGGCGGAGAG GGCGCGCAA CCGGCTMAGT AATTAGGTUG ACCGTGCTC  
5001 GGTTCGCCGA CTGGAAGCG GGCAGTGAGC GCACGCAAT TAATGTGAGT TACCTCACTC ATTAGGCACC CCAGGCTTTA CACTTATGC TTCCGGCTC  
CCAAAGGGCT GACCTTCGC CCGTCACTCG CGTTCGTTA ATTACACTCA ATGGAGTCAG TAATCCGTGG GGTCCGRAAT GTGAAATAGG AAGGCGGAG  
5101 TATGTTGTGT GGAATGTGA GCGGTAACA ATTCACACA GGAACAGCT ATGACCATGA TTACGAATTA A  
ATACAACACA CCTTAACT CCGCTATTGT TAAAGTGTGT CCTTTGTGGA TACTGGTACT AATGCTTAAT T  
>length: 5171
```

Figure 3-4

Figure 4
PSV.IPD
length: 5712 (circular)

1 TTCGACTCG CCCGACATTTG ATATTGACT AGATCGATC GACGCTGTG GAATGTGTGT CAGTTAGGTT GTGGAAAGTCC CCCAGGCTCC CCACGAGGCA
 AAGCTCGAGC GGGCTGTAAC TAATAACTGA TCTCAGCTAG CTCTCGACAC CTTACACACA GTCAATCCCA CACCTTTCCG GGGTCCGAGG GGTCTGTCCTG

101 GAAAGTATGCA AAGCATGTCAT CTCAAATTAGT CAGCAACCCAG GTGTGAAAG TCCCCAGGCT CCCAGCAGG CAGAAGTATG CAAGACATGC ATCTCAATTA
 CTTACATCGT TTGCTACGTA GAGTTAATCA GTCTGTGGTC CACACCTTTC AGGGTCCGA GGGTCTGATC GTCTTCATATG GTTTCTGATG TAGAGTTAAT

201 GTCAGCAAC CCATGTCCTCC CCCTACTCC GCGCCATCCG CCCTACTC CGCCCATTTT CCGCCCCATG GCTACACTAAT TTTTTTTTTAT
 CAGTCTTTGG TATCAGGGG GGGATTGAGG CCGGTAGGCG GGGATTGAG CCGGGTCAAG CCGGGTAAAG GCGGGGTAC CGACTGAATA AAAAAAATAA

301 TATGCAAGG CCGAGGCCGC CTCGGCCTCT GAGCTATTC ABAAGTAGT AGAGGCTTT TTTGGAGGCC TAGGCTTTTG CAAAAAGCTA GCTTATCCUG
 ATACGTCYCC GECTCGGGG GAGCCGGAGA CTGATPARG TCTTCATCAC TCCTCCGAAA AAACCTCCGG ATCGRAAAC GTTTTTCGAT CGAATPAGGUC

401 CCGGAAACG TGCATTGGA CCGGATPCC CCGTCCCAAG APTGACGTAA GTACCGCTTA TAGAGCGACT AGTCCACCAT GACCGAATAC AAGCCCALUC
 GGGCTTGGC ACGTAACCTT GCGCTAAG GGCACGGTTC TCACTGATTT CATGGCGAT ATCTCGCTGA TCAGGTGTA TCAGCTCATG TTTCCGGTCC

501 TCGGCTGCG CACCCGGAC GACGTCCTCC GGGCCGTACG CCGCTCTCC GCGCGTTCG CCGACTACCC CCGCACCGC CACACCGTAG ACCGAGALUJ
 ACCCGAGCG GTGGGGCTG CTGCAGGGG CCGGGCATGC GTGGAGCG CCGGTATGCG GCGGTCCGCG GTGTGGCATC TGGGCTTGGC

601 CCACATCGAG CCGGTACCG AACTGCAAGA ACTCTTCTC AGCGGCTCG GGTCCGACAT CCGCAAGTGT TGGTCCGCG ACBACUJUC CULUJTTUJLJ
 GGTPTAGCTC GCCCAGTGG TCGAGTCTT TGAGAAGGAG TCGCCGACG CCGAGTGTGA GCGGTCCAC ACCCAGCGCC TGTTCGCGG GCGCALUJLJ

701 GTCTGGACCA CCGCGAGAG GTCGAAGCG GGGGGGTGT TCGCCGAGAT CCGCCCGCGC ATGGCCGAGT TGAGCCGTTG CCGCTGCGC GUDALALANLJ
 CAGACTGTT GCGGCTCTC GCAGCTTCC CCGCCGACA AGCGGCTCTA GCGGGGCGG TACCGGCTCA ACTCGCCAG GCGCGACCGG CGLUJCTUJTG

801 AGATGGAAG CCTCCTGGG CCGCACCGG CCAAGAGCC CCGTGGTTC CTGGCCACC TGCGGCTCTC GCGCGACCAC CAGGGCAAGG GTCTGCGUJAG
 TCTACCTTCC GGAGACCGC GCGTGGCGG GGTTCCTCG GCGCACCAAG GACCGGTGGC AGCCGACAG CCGGCTGGTG GTCCCGTCC CAGACCCUJTC

901 CCGGCTGCTG CTCGCCGAG TGGAGGCGC CGAGCGGCC GGGTGGCCG CTTCTGGA GACTTCCG CCGCGAAC TCCTTTCTA GAGUJGUCJLJ
 GCGCAGCAC GAGGGCCTC ACTCCGCGG GCTCGCGCG CCCCAGGCG GAAAGACTT CTGGAGCGC GGGCCGTTG AGGGGAAGAT GCTTCCUJCG

1001 GCTTCCCG TCACGCGCGA GGTGAGTGC CCGAAGGACC GCGGACCTG GTGCATGACC CCGAAGCCCG GTGCCAACAT GGTTCGACCA TTTGAALTLCA
 CCGAAGTGGC AGTGGGGCT GCAGCTCAG GCTTCTCTG CCGGCTGGAC CACSTACTG GGGTTCGGC CAGGTTGTA CCAAGCTGGT AACTTACGCT

1101 TCGTCCCGT GTCCRAAAT ATGGGATG GAAAGAGG AGACCTACC TGCCCTCCG TCAGGAACG GTTCAAGTAC TTCCAAAGAA TACALALANLJ
 AGCAGCGCA CAGGTTTTA TACCCCTAAC GGTCTTCCG CCGGCTGGAC CCGGCTGGAC CACSTACTG GGGTTCGGC CAGGTTGTA CCAAGCTGGT AACTTACGCT

1201 CTCTTCACTG GAAGTAAAC AGAATCTGTT GATTATGGT AGRAAACC GTTCTCCAT TCCTGAGAG AATCGACCTT TAAAGACAG AATTAATATA
 GAGAGTAC CTTCCATTTG TCTTAGCCA CTATACCCA TCCTTTTGA CCAAGAGTA AGGACTCTTC TTAGTGGAA ATTTCTGTC TTAATTAAT

1301 GTTCTCAGTA GAAACTCAA AGAACACCA CGAGGACTC ATTTCTTGC CAAAAGTTG GATGATGCT TAAAGCTTAT TGAACAACCG GAATTTGATA
 CAAGAGTCAI CTCTTGAGTT TCTTGGTGGT GTCCTCGAG TAAAGRACG GTTTTCAAC CTACTACGGA ATTTCTGATA ACTTGTGCG CTTTACUJTT

1401 GTAAGTAGA CATGGTTGG ATAGTCGGAG GCATTTCTGT TTACAGGAA GCCATGAAAT AACCAAGGCA CTTTAGACTC TTTCTGACAA GUAATTAATATA
 CATTTCATCT GTACCAACC TATCAGCTC CGTCAAGACA AATGCTCCTT CCGTACTTAG TTGGTCCGCT GAACTCTGAG AAACACTUJTT CTTAGTATAJLJ

1501 GGAATTTGAA AGTGACAGT TTTTCCAGA AATGTTTG GGAATATA AACCTCTCCC AGAATACCCA GCGCTCTCT CTGAAJTLCA GAGUJGUAJLJ

Figure 4-1

CCTTAAACTT TCACCTGTGCA AAAGGGTCT TTRACTAAAC CCCTTTATAT TTGGAGAGGG TCTTATGGGT CCGCAGGAGA GACTCCAGGS' CCTCCCTTTT
 1601 GGCATCAAGT ATRAGTTTGA AGTCTACGAG AAGAAGACT AACGTTAAT GCTCCCTCC TCAGCTATG CATTTTTATA AGACCAATGG ACTTTTTCCTG
 CCGTAGTTCA TATTCAACT TCAGATGCTC TTCTTTCTGA TTGCATTTGA CGAGGGGAGG AITTCGATAC GTAAAAATAT TCTGGTACCC TGAANAACGAC
 *End DHER
 1701 GCTTAGATC CCTTGGCTT CATTAGAGG CAGCTACAAAT TATACATAA CCTTATGAT CATACACATA CGATTTAGG GACACTATAG AFAACATCCA
 CAAATCTAG GGAACCGRA GCAATCTGG CTCGATGTA ATTAGTATT GGAATACATA GTATGTAT GCTAAATCCA CTGTGATAT TATTCTAGCT
 1801 CTTTGCCTT CTCTCCACG GTGTCCACTC CCAGTCCAA CTGCACCTCG GTCTATCGA TTGAATCCC CGGGATCTT CTAGACTCGA CCFGCAGAA;
 GAAACGGAAA GAGAGGTGC CACAGTSHG GGTCCGGTT GAGCTGCAC CAGATAGCT AACTTAAAGG GCCCTTAGG GACTCTAGG GACCTCTTTC
 1901 CTTGATGGC GGCATGGCC CAATTTGTT AVTGCAGCT ATATGGTTA CAATTAAGC AATAGCATCA CAATTTTAC AATTAAGCA TTTTCTTAC
 GAAGTACCG GCGTACCGG GTTGAACAA TAAGTCCRA TATTCACAT GTTATTTCC TTATCTAGT GTTAAAGTG TTTTCTTCT AAAAAAAGT;
 2001 TGCATCTAG TTGTGTTG TCCAACTCA TCAATGTATC TTATCATGC TGGATCGATC GGAATTTAT TCGCCGAGC ACCCTGGCT GAAATAACT
 ACSTAGATC AACACCAAC AGSTTTGAGT AGTTACTATG AATAGTACAG ACCTAGCTAG CCTTAAATTA AGCCCGCTG TGGTACCCGA CTTTATTTGA
 2101 CTGAAGAGG AACTTGGTTA GGTACCTTCT GAGCGGAAA GAACCACTG TGGATGTTG GTGTGAAAAG TCCCAGGCT CCCGAGCAGG
 GACTTTCTCC TTGACCAAT CCATGGAAGA CTCCGCTTCT CTGCTGAGC ACCTTACACA CAGTCAATCC CACACCTTTC AGGGCTCGA GGGCTCTTCC
 2201 CAGAAATAG CAAGCATGC ATCTCAATTA CTCAGCAACD AGTGTGGAA AGTCCCAGG CTCGCCAGCA GGCAGAGTA TGCRAAGCT GCATCTCAAT
 GTCTTCATAC GTTCTGACG TAGATTAAT CAGTCTGTGG TCCACACTT TCAGGGTCC GAGGGTCT CCGTCTTCAT ACGTTTCTGA CGTAGAGTTA
 2301 TAGTCAGAA CCATAGTCC GCGCTACT CCGCCATCC CCGCCCTAAC TCCGCCAGT TCCGCCCAT CTCCGCCCA TGGCTGACTA ATTTTTTTTTTA
 ATCAGTCTGT GGTATCAGG GGGGATTA GCGGGTAGG GCGGGATG AGGGGTCA AGCGGGTAA GAGGGGGGT ACCGACTGAT TAAAAAATA
 2401 TTTATCAGA GCGCAGGC CCTCGGCT CTGAGTATT CCAGAATAG TGAGAGGCT TTTTGGAG CTTAGCTTT TCCAAAAC TCTTTACTTCC
 AATAGCTCT CCGGCTCGG GAGAGCGGA GACTCGATA GCTCTTATC ACTCTCCGA ABAACCTCC GATCCGAAA ACGTTTTTC ACAAATGAGG
 2501 AGCGGGCT TAAATAGG GGCATTTA AATCTGCAG SPAACAGCT GGACTGACC CCGTGTAC CCGTGAACCG TGCAGACT GACCTTTTGG GACCCGAAAT
 TCGCCGGGA ATTAATCCG CCGGTAAT TTAGGACGTC CATTGTGAA CCGTGAACCG CAGCAATG TGCAGACT GACCTTTTGG GACCCGAAAT
 2601 CCCACTTAA TCGCCTTGA GCATCCCC CCFCCGAG CTGGGTAAT AGCAAGAG CCGCACCGA TCGCCCTTCC CAGCAGTTG GTAGCTTAAA
 GGGTTGANT AGCGAAGT CTTGAGGG GGAAGGCT GACCCATTA TCGCTTCTCC GCGCTGGCT AGCGGAAAG GTTGTCAACG CATCGCACT
 2701 TGGCATGG CCGTATGC GTATTTCT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT CTTAGCCTT
 ACCGCTTACC GCGACTACG CCAARAAGA GGAATGCTTA GACAGGCTT AAGTGTGC CTAGCTTCC GCGCTGGCT AGCGGAAAG GTTGTCAACG CATCGCACT
 2801 CATTRAGC GCGGGTGT GTGTTAGC GCAGGCTAC CCGTACTT CCGATGMA CCGTCCGCG TCGCCCGG CCGCTTCCCT TCTTCTCCCT CTTTTCTCC
 GTAATTCGG CCGCCACAC CACCAATGG CCGTACTG CCGATGMA CCGTCCGCG CCGTCCGCG TCGCCCGG CCGCTTCCCT TCTTCTCCCT CTTTTCTCC
 2901 CAGCTTCC GCTTCCCG GCAAGTCT AATCGGGG CTTCCCTTAT GATGCTTTA CCGCACCTCG ACCCAAAA ACTTTGATTTG
 GTCAAGCG CCGAAGGG CAGTGGGA TTTAGCCCC GAGGAATC CCAAGCTTA ATCAGGANT CCGTGGAG TGGGTTTTT TGAATAAA
 3001 GGTGATGTT CAGTATGG GCAATGCC TGTATAGCG TTTTCCCG TTTAGCTTG GAGTCCAGT TCTTATAG TGCATCTTG TCCAAACTU
 CCACTACC AA GTGCATCACC CCGTACCGG ACTATCTGC AAAAGCGG AACTGCRAC CTAGGTGCA AGAATATC ACTGAGAAC AAGCTTTGAC
 3101 GACACACT CAACCTATC TCGGCTATT CTTTGTATT ATAGGANT TTGCGGTTT GCGCTATTT GTTAAAAAT CAGCTGATTT AAAAAAAT
 CTTGTTGTA GTTGGATG AGCCGATAA GAAACTAAA TATTCCTAA TATTCCTAA TATTCCTAA TATTCCTAA TATTCCTAA TATTCCTAA
 3201 TAACCGGAT TTTAACAAA TATTAAGTT TACAATTTA TGGTCACTC TCAGTACAT CTGCTCTGAT GCGCATAGT TAAAGCAACT CACTATCTCC
 ATTGGCTTA AATTTGTTTT ATAAATGCA ATGTTAAAT ACCACGTGAG AGTCAATGTA GACGACTA CCGCGTATCA ATTCGTTTGA GCGCTATGCT

Figure 4-2

3301 TAGGTGACTG GGTCTATGGCT GCGCCCGGAC ACCGGCCAC ACCCCCTGAC GGCCTTGCT GCTCCCGGCA TCCGCTHACA GACAAGCTGT
ATGCACTGAC CCAGTACCGA CCGGGGGCTG TGGGGGACTG CCGGACAGA CAGGGCCGT AGGGAAATGT CTGTTCGACA

3401 GACCGTCC GGGAGTGA TGTGTGAG GTTTTCGCG TCATCACGA AAGCGGGAG GCAATATCT TGAAGACGAA AGGCCTCCT GATACGGCTA
CTGGCAGAG CCTCGAGCT ACACAGTCT CAAAAGTGG ACTAGTGGT TTGCGCGCT CCFATAGA ACTTCTGCT TCCCGGACA CTATCCGANT

3501 TTTTTRAG TTAATGTCAT GATATAATG GTTCTTAGA CBTGAGTGG CACTTTTCGG GGAATGTGC GGGAAACCC TATTTGTATA TTTTTCYAAA
AAAAATATCC AATTACAGTA CTATTATTAC CAAAAGATCT GCAGTCCACC GTGAABAGCC CCTTTACAGC CCGCTTGGGG ATNAACAAT AAAAAGATTT

3601 TACATTCAAA TATGTATCCG CTCATGAGAC ATAAACCCCTG ATAAATGCTT CAATATATTT GBAABAGGA GATATGACT ATTCACATTT TCCGTGTCCG
ATGTAAGTTT ATACATAGCC GAGTACTCTG TTAATGGGAC TATTTACGAA GTTATATATA CTTTTCCCTT CTCATACTCA TAAGTTGTAA AGGCACAGCC

3701 CCTATTCC TTTTTCGG CATTTGCC TCTGTGTTTT GCTCACCCAG AAGCCTGCT GBAAGTAAA GATGCTGAG ATCAGTTCGG TGCACGACTG
GGATYAGGG AAAAARCCG GTAAACGGA AGCAAAAHA CAGTGGGTG TTGCGRACA CTTTCAATTT CTACGACTTC TAGTCAACCC ACBTGCTCAC

3801 GGTACATCG AACTGGATCT CAACAGCGGT AGATCCTTG AGATTTTCG CCGGAAAGAA CTTTTCCAA CTTTTCCAA TCTGTGACTG TCTGTGACTG
CCATGTAGC TTGACCTAGA GTTGTCCCA TTCTAGGAA TCTAAAGC CCGCTTCTT GCAAAAGGT ACTACTCTG GAAATTTCA GACGATACAC

3901 GCGCGTAT ATCCCGTGT GAGCGGGC AAGGCAACT CGTCCGCGC ATACACTATT CTCAGATGA CTTGTTGAG TACTCACCG TCACAGAAAA
CGGCAATA TAGGCACCTA CTGGGGCCG TTCTGCTGA CCGAGCGCG TATGTATA GAGTCTTACT GAACCACTC ATGAGTGGT AGTGTCTTTT

4001 GCATCTAG GATGCATGA CAGTAAGAGA ATTATGAGT GCTGCAATA CCATGATGA TAACACTGG GCCACTTAC TCTGTGACAAC GATCCGNGA
CGTGAATCC CTACGCTACT GTCAATCTCT TAATAGCTCA CGAGGTAFT GGTACTACT ATTTGAGCC CGTGTGAATG AAGACTGTG CTAGCCCTCT

4101 CCGRAGGAG TARGCTTT TTTGCACAAC ATGGGGATC ATGTAACCTG CTTTGTCTG TGGAAACCG AGCTGAATA RGCATACCA AACACAAAC
GGTCTCCG ATTTGCGAAA AAGGTGTTG TACCCCTTAG TACATGAGC. GAACTAGCA ACCCTGGCC TGGACTTACT TCGGTATGTT TTGCTGCTG

4201 GTGRCACAC GATCCGCA GCAATGSCRA CAACGTGCG CAATATTA ACTGGCAAC TACTACTCT AGCTTCCCG CAACAATAA TAGACTGAT
CACTGTGTTG CTGCGTCTG CTTACCGTT GTTGCACCC GTTTGATAT TGACCCGTTG ATGAATGAGA TCGAAGGCC GTTGTAAAT ATCTGACCTA

4301 GAGGGGGAT AAGTTGCRG GAGCACTTCT GCGCTCGCC CTCCGGCTG GCTGGTTAT TCGTGATRAA TCTGAGCGG GTGAGCGTG GTCTCCGCT
CCTCCGCTA TTTCAAGTC CTGGTGAAGA CCGRAGCGG GAAGCCGAC CGACAAATA AGCACTATTT AGACTGCGC CACTGCGCC CAGAGCGCCA

4401 ATCAITGAG CACTGGGCC AGATGGTAA GCGTCCGTA TCGTAGTAT CTACAGAGC GGGAGTCCAG CAACTATGGA TGAACGAAAT AGACAGATCG
TAGTRAGTC GTGACCCCG TCTACCAATC GGGGGGAT AGCTTCAATA GATGTCTGC CCTCAGTCC. GTTGTACTT ACTTGTCTTA TCTGTCTAGC

4501 CTGRTAGG TGCCCTACTG ATTAAGCATT GGTAACTGTC AGACCAAGTT TACTATATA TACTTTAGAT TCAATTAHA CTTCAATTTT AATTTAAAG
GACTCTTCC ACGGATGRC TRATTTGTA CCAITGACAG TCTGTTCAA ATGAGTATAT ATCAATCTA ACTAAATTTT GAAGTAAAA TTAATTTTC

4601 GATCTAGTG AAGATCCTTT TTGRTAATCT CATGACCAAA ATCCCTTAACT GTGRTTTTC GTTCCACTGA CCGTCAGACC CCGTCAAAA GATCAAAAGG
CTAGTCCAC TTCTAGGAA ACTATTTAGA GTACTGTTT TAGGRATG CACTCAAAAG CAGGTGACT CCGACTGCG GGCATCTTTT CTAGTTTCTT

4701 TCTTCTGAG ATCCTTTTTT TCTGCGGTA ATCTGCTCT ATCAACAAA AAACCAACG CTACAGCGG TGTTTGTTT GCGGAAACA GAGTACCAA
AGAGACTC TAGBARAAA AGACGGCAT TAGCAACA CTAGGACTG TTTTGTGGC GATGCTGCC ACCAAACAAA CCGCTTACTT CTGATGTTT

4801 CTCTTTTCC GAGGTAAT GSCCTCAGCA GAGCGAGAT ACCAAATACT GTCTTCTAG TGTAGCCGTA GTTAGGCCAC CACTTCAAGA ACTCTGTAUC
GAGBARAAG CTTCATTTGA CCGAATCTGT CTGCGCTA TGTTTTATGA CAGRAGATC ACATCCGAT CATCCGGTGT GTGAAGTTCT TGAGACATCG

4901 ACCGCTACA TACCTGCTC TGTATCTCT GTTACCAGTG GCTGTGCCA GTGGGATAA CTGCTCTT ACCGTTGCTT ACCGTTGCTT ACTCAAGAGC ATCTGTAAT
TGGGGATGT ATGAGGGAG ACGATTTAGA CAATGCTCAC CGAGGCGGT CACCGTATT CAGCAGAA TGCCCAAC TGAGTTCTGCT TATCAATGUC

5001 GATAGGGCC AGCGGTGCGG CTGAACCGG GGTTCGTCA CACAGCCCG CTTGAGCGA CCGACTTAC CCGAAGTACAG CTTGACATTT

Figure 4-3

CTATTCCGGC TCGCCAGCCC GACTTGCCCC CCARGCAGGT GTGTGGGGTC GAACCTCGCT TSCITGGATGT GGCITGACTC TATGGATGTC GCACCTCGTAA
 5101 GAGAAAGCC CACGCTTCCC GAAGGGAGAA AGCGGGACAG GTATCCGGTA AGCGGCAGG TCGGAACACAGG AGAGGGCAGG AGGGAGCTTC CAGGGGGAAA
 CTCTTTCCGG GTGCGAAGGG CTTCCCTCTT TCCGCCCTGT CATAGGCCAT TCGCGCTCCC AGCCTTGTCC TCTCGGCTGC TCCCTCGAAG GTCCCCCTTT
 5201 CGCCTGGTAT CTTTATAGTC CTGTCCGGTT TCGCCACCTC TGACTTGACG GTCGATTTTT GTGATGCTCG TCAGGGGGGC GGAGCCTATG GAAAAACCCC
 GCGGACATA GAAATATCAG GACAGCCCA AGCGTGGAG ACTGAACCTG CAGCTAAAAA CACTACGAGC AGTCCCCCG CCTCGGATAC CTTTTTCGGG
 5301 AGCAACGGC CTTTTTACG GTTCTGGCC TTTTGTGGC CTTTTGCTCA CATGTTCTTT CCTGGTTAT CCCCTGATTC TGTGGATAAC CGTATTACCG
 TCGTTGGCC GGAAAAATGC CAAGGACCG ANACGACCG GAAACGAGT GTACAAGAAA GGACGCRATA GGGGACTAAG ACACCTATTG GCATATAATGGC
 5401 CCTTTGAGT AGCTGATACC GCTGCGGCA GCGGACGAC CGAGCGCAGC GAGTCAGTGA CCGGAGAAGC GGAAGAGCC CCAATFACGA AACCGOCT
 GGAACCTCAC TCGACTATGG CGAGCGGGT CGGCTTGTG GCTCGGCTG CTCAGTCACT CGCTCCTTGG CCTTCTCGG GGTATFSGT TTGGCGGAGA
 5501 CCCCAGCGT TGGCCGATC ATTAATCCAG CTGGCAGCAG AGGTTCCCG ACTGGAAGC GGGCAGTGA GCGAAGCAA TTAATGTAG TPACTTCACT
 GGGGGGCA ACCGGCTAAG TAATTAGTC GACCGTGTG TCCRAAGGC TGACCTTTC CCGCTCAC TC GGTTCGTT RATTACACTC AATGGAGTGA
 5601 CATRGGCAC CCCAGGCTTT AACTTTATG CTTCGGCTC GTATGTTGTG TGGATTTG AGCGGATAC AATTTACAC AGGAAACAGC TATGACCATC
 GTARTCGGTG GGGTCCGAAA TGTGAAATAC GAAGCCCGAG CATACACAC ACCTTAACAC TCGCCTATTG TTAAGTGTG TCCTTTGTTCG APTACTGATAC
 5701 AATFAGAAAT AA
 TAATGCTTAA TT

>length: 5712

Figure 4-4

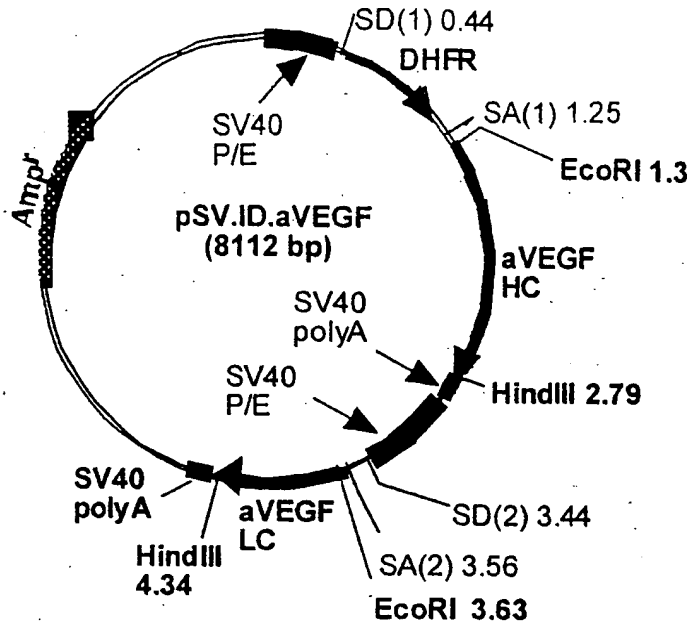


Figure 5, pSV.ID.aVEGF control plasmid

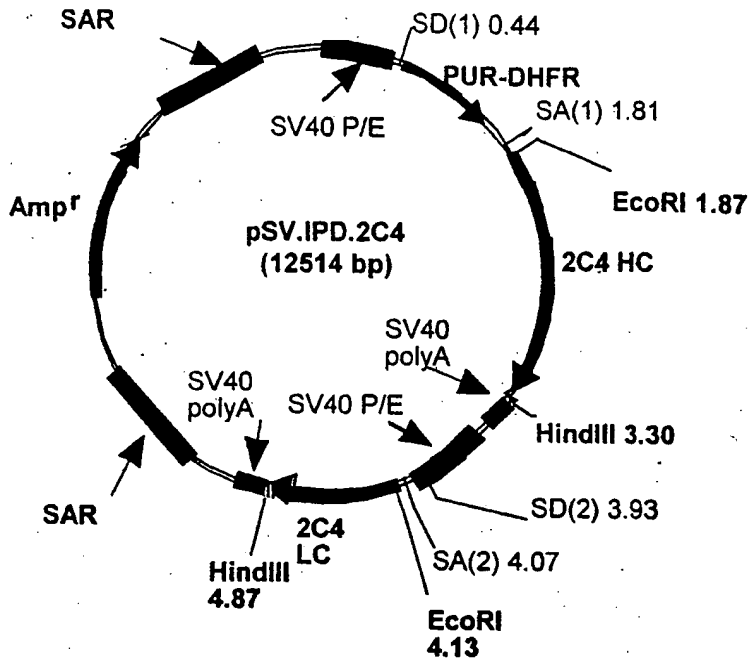


Figure 6. pSV.IPD.2C4

Figure 7
PSV.IPD.2C4
length: 12514 (circular)

1 TTCCAGCTCC CCCGACATTC ATTATTGACT AGATCGGATC GACAGCTGTG GAATGTGTGT CAGTTAGGTT GTGRRAGTC CCCAGGGCTCC CCAGCAGGCA
AAGCTCGAGC GGGCTGTATC TAATAACTGA TCTCAGCTAG TCTGCGACAC CTACACACA GTCAATCCCA CACTTTTCAG GGTCCGAGG GGTCTGCGT

101 GAAGTATGCA AAGCATGAT CTCAATTAGT CAGCARCCAG STGTGGAAG TCCCCAGGCT CCCACACAGG CAGAAGTATG CAAGAATGTC ATCTCAATTA
CTTCATACGT TTCTACGTA GAGTTAATCA GTCTGTGTC CACACTTTC CACACTTTC AGGGTCCGA GGGGTCTGCC GTTTCATAC GTTTCGTAAG TAGAGTTAAT

201 GTACGACAC ATAGTCCCGC CCGTAATCC GCGCATCCG CCGTAACATC CGCCAGTTC CGCCCATTTCT CGCCCATTTCT CGCCCATTTCT TTTTTTTTAT
CACTCGTTGG TATCAGGGCG GGGATTGAG GGGTAGGC GGGATTGAG GGGGTCAAG GGGGTAAAG GCGGGGTAC CGACTGATTA AAAAAAATA

301 TATGCGAGG CCGAGCGCG CTGCGCTCT GAGCTATTCC AGRAGTAGT AGGAGCTTT TTGGAGGC TAGGCTTTG CAAAAAGCTA GCTTATCCGG
ATAGCTTCC GGTCCGCGC GAGCGGAGA CTCGATAAG TCTTCATCAC TCTCCGAAA AACCTCCG ATCCGAAAAC GTTTTCGAT CGAATAGGCG

401 CCGGGAACGG TGCATTGAA CGCGATTCC CCGTCCNAG ATRAGCTAA GTACCGCTA TAGAGCGACT AGTCCACCAT GACCGAGTAC AAGCCACCG
GGCCCTGCC ACCTAACCTT GCGCTAAG GGCAGGTTT TCACTGAT CATGCGGAT ATCTGCTGA TCAGGTGTA CTGGCTCATG TTCCGGTGC

501 TCGCCCTCC CACCCGGAC GACTCCCGC GGGCCGTAG CACCTCCG CCGGCTTCG CCGACTACC CGCCACGCG CACACGCTAG ACCCGACCG
ACCCGGAGCG GTGGCGCTG CTGCAAGCG CCGGCATGC GTGGAGCG GGGTATGG GGGTATGG GGGTATGG GGGTATGG GGGTATGG GGGTATGG

601 CCACATCCG CCGGTACCG AGTGCAGA ACTTCTCT ACGCCCTCC GGTTCGACT CGGCAAGTG TGGTCCGGC ACGAGCGCG CCGGTTCCG
GGTGTAGCT GCGCATGGC TCGAGCTTCT TGAAGAGG TCGCGGAG TCGCGGAGT GCGGTCCAC ACCAGCGCC TCGTCCGCG GCGTCCGCG

701 GTCTGACCA CGCCGGAGG GGTGAGCG GGGCGGTGT TCGCCGAGT CGGCCCGC ATGCGCGAGT TGAGCGGTC CCGGCTGCC CGCAGCAAC
CAGACTGTG GCGCCCTC GCAGCTTCC CCGCCACA AGCGCTCA GCGGCGCG TACCGCTCA ACTCGCAAG CCGCAGCG CCGGCTGCC

801 AGATGGAAG CCTCTGGC CCGCACCGC CCAAGGAGC CGGTGTTT CTGGCCATC TGGCCGTC CCGCACAC CAGGCAAG GTCTGCGCAG
TCTACCTCC GGAGGACCG GCGGTGCCG GTTCTCCG GCGACCAAG GACCGGTGG AGCGCGAG CCGGCTGGT GTCCCGTTC CAGACCTTC

901 CCGCTCTG CTCCCGGAG TGGAGGCG CGAGCGCG CCGGTCGCG CCGGTCGCG GACTCCCG CCGCGAAC TCCCTTCTA CGAGCGCT
CGCGCAGCAC GAGGGCTC ACCTCGCG GTCGCGCG CCGCGCGG GAAAGACCT CTGAGGCG GGGCTTGG AGGGAAGT GCTCCCGAG

1001 GCTTCACG TCACCGCCG CGTGAGTC CCGAGGAC GCGGACTG GTCAATGAC CGCAAGCCG GTGCAACAT GTTCGACA TTGAAGTCT
CGAAGTGC AGTGGGGT GCGCTCAG GCTTCTCG GCGCTTCC GCGCTTCC GCGCTTCC CCAAGCTGT AACTGACT

1101 TCGTCCCG GTCCCAAT ATGGGATT GCAAGACGG AGACTTACC TCGCTCCG CCGGCTGG CAGTACTG GGTTCGGG CAGGTTGTA CCAAGCTGT AACTGACT
AGCAGCGCA CAGGTTTA TACCCCTAAC CGTCTTGC TCTGATGG AGGGAGCG ATCTCTGG CAGTACTG GGTTCGGG CAGGTTGTA CCAAGCTGT AACTGACT

1201 CTCTTCAGT GAAGTAAAC AGAATCTGT GATTATGGT AGMAACCT GGTCTCCAT TCGTGAAG AATCGACTT TAAAGACAG AATTAATATA
GAGAAGTCACTTCCATTG TCTTAGACCA CTAAATACCA TCCCTTGA CCAAGAGTA AGGACTCTT TTAGTGGAA ATTTCTGTC TTAATATA

1301 GTTCTACTA GAGACTCAA AGACACCA CGAGGCTC ATTTCTTCC CAAAATTG GATGATCT TAAAGACTT TGAACACCG GAATTCGCA
CAAGACTCAT CTCTGAGTT TCTTGGTT GCTCTCCG GCTCTCCG TAAAGAGG GTTTTCAAC CTACTAGGA ATTTCTGTA ACTTCTGTC CTTAACCGT

1401 GTAAGTAGA CATGTTTGG ATAGTCGG GAGTCTGT TTACAGGA CCAATGAATC ARCCAGGCA CTTAGACAC TTTGTCACAA GATCATAA
CACTTCTCT GTACCAAAC TATCAGCTC GGTCAAGACA AATGCTCTT CCGTACTTAG TTGTCGCTT GAACTCTG AACTCTGTT CTTAGTACT

1501 GGAATTTGAA AGTACACGT TTTCCCGA AATTGATT GGAATATA AACCTCTCC AGAATACCA GCGTCTCT CTGAGGTCCA CGAGGAAAA

Figure 7-1

CCTTAAACIT TCACTGTGCA AAAAGGTCT TTAACATAAC CCCTTTATAT TTGGAGAGGG TCTTATGGGT CCGCAGGAGA GACTCCAGGT CCTCCTTTT

1601 GGCATCAAGT ATAAGTTTGA AGTCTAGGAG AGAARAGACT AAGGTTAACT GCTCCCTCC TAAAGCTATG CATTTTTATA AGACCATGG ACTTTTGGCTG
 CCGTAGTTCA TATTCHAACT TCAGATGCTC TTCTTTCTGA TTGCAATTGA CGAGGGGNG ATTTGCGATAC GTAARAATAT TCTGGTACCC TGAARAACGAC

1701 GCTTTAGATC CCGTTGGCTT CGTTAGAACG CAGTCAAT TAATACATAA CCTTATGTAT CATACRCATA CGATTTAGGT GACACTATAG AATACATCC
 CGAATCTAG GGGAAACCGAA GCAATCTGCG CTCGATGTTA ATATGTAATY GBRATACATA GTATGTGTAT GCTAATAATCCA CTGTGATATC TTATTTGTAAG

1801 ACTTTGCCIT TCTTCCACA GGTGTCCAAT CCGAGTCCA ACTGCACCTC GGTTCATCG ATTGAATTC ACCATGGGAT GGTCTGTAT CATCCTTTT
 TGRAACGAA AGAGAGGTET CCACAGGTA GGTTCAGGT TGCCTGAGG CCAAGATAGC TAACHTAAGG TGSTACCTA CCAATACATA GTAGGAAAA

1901 CTAGTAGCAA CTGCAACTGG AGTACATTCG GAAGTTCCAG TGGTGGATC TGGCGTGGC CTGGTGCAG CAGGGGCTC ACTCCGTTG TCCTGTGCRAG
 GATCATCTGT GACGTTGACC TCATGTAAGT CTTCAAGTUG ACCACCTCAG ACCGCCACCG GACCACCTCG GTCCCCGAG TGAGGCAAAAC AGGACACCTC

2001 CTTCTGGCTT CACCTTACC GACTATACCA TGGACTGGGT CCGGTGAAGG GCTTGAATG GTTTGCAGAT GTTAATCCTA ACAGTGGCGG
 GAARCCGAA GTGGAAATGG CTGATATGTT ACCTGROCCA GGCACCTCCG GCGCCATTC CCGACCTTAC CCAACCTCTA CAATTAGAT TGTACCCGCC

2101 CTTCTATCTAT AACGAGGCT TCRAGGGCGG TTACTCTG AGTGTGACA GATTAATAA CACTTATAC CTGCAGATGA ACAGCTGCG TGCTGAGGAC
 GAGATGATA TTGTTGCGGA AGTTCCCGG AAGTGERGAC TCACAACTCT CTAGATTTT CTGTAATATG GAGTCTACT TGTCCGACCG ACGACTCTCG

2201 ACTGCGTCT ATTAFTGTG TCGTAACCTG GACCTCTT TCTACTTGA CTACTGGGT CAAGGAAACC TGSTCACCGT CTCTCGCC TCCACCRAGG
 TGACGGCAGA TAATAACAG AGCATTTGAC CCTGGGAGNA AGATGAACT GATGACCCCA GTTCTTTGG ACCAGTGGCA GAGGACCGG AGSTGGTTCC

2301 GCCATCGGT TTCCCCCTG GCACCTCT CCAAGAGCAC CTCTGGGGC ACAGGGGCC TGGGTGCTT GGTTCARAGG TACTTCCCG AACGGGTGAC
 CCGGTAGCCA GAAGGGGAC GGTGGGAGG GATTTCTGNS GAGACCCCGG TGTCCCGGG ACCCGACGGA CCGATTCCTG ATGAAGGGG TTGGCCACTG

2401 GGTGTCTGG AACTCAGGCG CCCTGACCAG CCGGTGCAC ACCTTCCCG CTGTCTTACA GTCTACTCC CTAGTCTCC TACGACAGCT GGTGTGAGG
 CCACAGCAC TTGATCCCG GGGACTGCT GCGGACAGTG TGGAGGGCC GACAGTGT CAGGATCTT GAGATGAGG GATGCTGCGA CCACTGTGAC

2501 CCCTTAGCA GCTTGGCAC CCAGACCTAC ATCTGCAAG TGRATCABA CCCCAGCAC ACCRAGGTGG ACMAAAGT TGAGCCCAA TCTTTGTGACA
 GGGAGATCGT CRRACCCGT GGTCTGGAG TAGAGTTGT ACTTAGTGT CCGGTCTGG TGGTTCCACC TGTCTTCA ACTCGGTTT AGAARCACTGT

2601 AACTCACAC ATGCCACCG TGCCAGCAC CTGACTCTT GGGGGACCG TCAGTCTTCC TCTTCCCGC AAAACCCAA GACACCTCA TGATCTCCC
 TTTGAGTGTG TAGGGTGGC ACGGTCTGT GACTTGAAGA CCCCCCTGC AGTCAGAAGG AGAAGGGGG TTTTGGGTTT CTGTGGGAGT ACTAGAGGGC

2701 GACCCCTGAG CTCACATGCG TGGTGGTGA CGTGGCCAC GAAGACCTG AGGTCAAGTT CACTGGTAC GTGGACGGG TGGAGGTGCA TAATGCCAAG
 CTGGGACTC CAGTGTAGC ACCACCACT GCACTCGTG CTCTGGGAC TCCAGTTCAA GTTGACCATG CACTGCGCG ACCTCCAGT ATTACGGTTC

2801 ACRAGCCCG GGGAGGACA GTACAACAG ACCTACCGG TGTTCAGGT CCTCACCGT CTGCACGAG ACTGGCTGA TGGCAAGGAG TACAAGTGA
 TGTTCGGCG CCTCCTCTGT CATGTTGCG TGCATGGCC ACCAGTCCA GGAGTGGCAG GAGTGTTC ACCGTTCTC ATGTTCACTG

2901 AGTCTCAA CRAAGCCTC CCRAGCCCA TCGAGAAC CACTCCAAA GCCAAGGGC AGCCCCGAG ACCACAGTG TACACCTCG CCCCATCCCG
 TCCAGAGTT GTTTCGGGAG GGTGGGGGT AGTCTTTTG GTAGAGTTT CCGTTTCCCG TCGGGTCT TGTGTCCAC ATGTGGGACG GGGTAGGCG

3001 GGAAGAGATG ACCAAGAAC AGTCTGAGCT GACTGCTG GTCRAGGCT TCTATCCAG CGACATGCC GTGAGTGGG AGACAATGG GCAGCCGGAG
 CCTTCTCTAC TGGTCTTGG TCCAGTGGGA CTGGACGGAC CAGTTTCCGA AGATGGGCT GCTGTAGCG CACCTCAC TCTGTTACC CGTCCGGCTC

3101 AACACTACA AGACACGCG TCCCTGCTG GACTCCAGC GCTCTTCTT CCTCTACAG AGCTCACCG TGGACAGAG CAGTGGCG CAGGGGAACG
 TTGTTGATGT TCTGGTGGG AGGCACGAC CTGAGGCTG CTGAGGAA GGAGATGCTG TTCCAGTGGC ACCTGTTCTC GTCACCGTC GTCCCTTTC

3201 TCTTCTCTAG CTCGGTGTG CATGAGCTC TGCACACCA CTACAGCAG AAGAGCTCT CCTGTCTCC GGTAAATGA GTCCGAGCG CTTAGAGTCC
 AGAAGATGAC GAGGACTAC GACTCCGAG AGTGTGTTGAT GATGTGGCTC TTCTCGGAGA GGCACAGAG CCAATTTACT CACGCTCCG GATCTCAGC

Figure 7-2

3301 ACCTGCAGAA GCTTCGATGG CCGCATGGC CCAACTGGT TATTCRAGT TATTAAGG CAATAGCAT ACAAAATTC ACAAATAAGC
 TGGACGCTT CGAAGCTACC GCGGATACC GGTGACAA ATRACGCGA ATATACCA TGTATTTC GTTATCGTAG TGTATAAGT GTTATTTC

3401 ATTTTTTCA CTGCATCTA GTGTGTTTT GTCCAACTC ATCAATGAT CTTATCATGT CTGGATCGG AATTAATCG GCGAGCACC ATGGCCTGAA
 TAAAAAAGT GACGTRAGT CAACACAAA CAGTTTTGAG TAGTTACATA GAATGATACA GACCTAGGCC TTATTAAGC GCGTCTGTGG TACCCGACTT

3501 ATACCTCTG AAAGAGAAC TTGTTAGT ACCTTCTG GGGAAAGAA CCAGCTGAG AATGTGTCT AGTATGAGT TGGAAAGTCC CCAGGCTCCC
 TATTGGAGC TTTCTCCTG AACCAATCCA TGGAAAGTCC GCGCTTCTT GGTGCACACC TTRACACACAG TCAATCCACAC ACCTTTCCAG GGTCCGAGGG

3601 CAGCAGCAG AGTATGCAA AGCATGCATC TCAATGATC AGCAACRAG TGTGGAAGT CCCAGGGTC CCRAGCAGC AGAAGTATCC AAAGCATGCA
 GTCGTCGTC TTCATACGTT TCGTACCTAG AGTTAATCAG TCGTTGTCTC ACACCTTCA GGGTCCGAG GGGTCTCCG TCTTCATACG TTTCTGACGT

3701 TCTCAATTAG TCAGCAACCA TAGTCCGCG CCTAACCTCC CCAATCCCG CCGCAATTC GCCCCCATGG CTGACTAAAT
 AGAGTTAATC AGTCGTTGTT ATCAGGGCGG GHTTGGAG GGGTATGAG GGGTCAAGG CCGGTAAGAG CCGGGTACC GACTGATTA

3801 TTTTTTATTT ATGCRAGGC CGAGGCGGCC TCGGCTCTG AGCTATTCCA GAAGTAGTGA GGAGGCTTTT TTGGAGGACT AGGTTTTTGC AAAAAGCTAG
 AAAAATARA TAGTCTCCG GCTCCGGCGG AGCGGAGAC TCGATRAGT CTTCACTACT CCTCCGAAA AACCTCCTGA TCCGAAAACG TTTTTCTGATC

3901 CTTATCCGGC CCGGAACGGT GCATTGGAAC GCGGATCCC CAGTCCAAAG TACCGCTAT AGAGTCTATA GCGCCACCCC CTTCGCTTCG
 GAATAGCCG GCGCTTGCCA CGTAACCTTG CCGCTAAGG GCAGGTTCT CAGTCCATTC ATGGCGGATA TCTCAGATAT CCGGTGCGG GAACCCGAAG

4001 TTAGAAGCCG GCTACAAATTA ATACATAAC TTTTGGATCG ATCTACTGA CACTGACATC CACTTTTTCT TTTTCTCCAC AGGTCTCCAC TCCCAGGTCC
 AATCTTCCG CGATGTTAAT TATGTTATTT ANAACCTTAG TAGGATGACT GTGACTCTAG GTGAAAAGA AAAAGAGGTG TCCACAGGTG AGGGTCCAGG

4101 AACTGCACCT CGGTTCCGGA AGCTAGTGT GCGTGAATCG ATTGAATTC ACCATGGAT GGTCAATGAT CATCTTTTT CTAGTAGCAA CTGCAACJGG
 TTGACGTGGA GCCAAGCGCT TCGATCGAAC CCGAGGTAGC TAACTAAGG TGTATCCCTA CCACTACATA GTAGAAAAA GATCATCGTT GACGTTGACC

4201 AGTACATTCA GATATCCGGA TGACCCAGTC CCGAGTCC CTGTCGCTT CTGTCGCGCT CTGTCGCGA TAGGCTCAC ATCACCTCA AGGCCAGTCA GGATGCTCT
 TCAATGAACT CTATAGTCT ACTGGTCCAG GGGTCCGAG GACAGCCGA GACACCCCT ATCCCGGTGG TAGTGGACT TCCGCTCAGT CCTACACAGA

4301 ATTGGTCTG CCTGGTATCA ACAGAACCA GGAAGCTC CGAACTACT GATTTACTCG GTCCTTACC GATACACTGG AGTCCCTTCT CCGTCTCTCTG
 TRACCACAGC GGACCATAGT TGTCTTTGGT CCTTTTCCAG CTTATGATG CTAATGAGC CGAAGATGG CTATGTGACC TCAGGGAAGA GCGAAGAGAC

4401 GATCCGGTTC TGGACGGAT TTCACTCTGA CCATCGCAG TCTGACGCA GAAGACTTCG CAACTTATTA CTGTCAACAA TATTATATTT ATCCTTACAC
 CTAGGCCAAG ACCCTGCCA AGTGAAGT GGTAGTCTG AGAGTGGT CTTCGAAAG GTTGAAATAT GACAGTTGTT ATATATATA TAGGAATGTC

4501 GTTTGGCAG GGTACCRAAG TGGAGATCRA ACGAATCTG GCTGCACCAT CTGTCTTCAT CTTCGCGCA TCTGATGAGC AGTTGAAATC TGGAACTGCT
 CAACCTGTC CCATGGTTC ACCTCTAGT TGCTTGACAC CGAGTGGTA GACAGAATA GAAGGCGGT AGACTACTCG TCAACTTTAG ACCTTGACGA

4601 TCTGTTGTT GCTGCTGTA TRACTTCTAT CCGAGAGG CCAAGTACA GTGGAAGGTG GATTAACGCC TCCANTCGG TAACCTCCAG GAGAGTCTCA
 AGACAACACA CGGACGACTT ATTTAAGATA GGTCTCTCC GFTTCTATGT CACTTCCAC CTATTCGCGG AGGTTAGCCC ATGAGGCTCTCTCACAGT

4701 CAGAGCAGGA CAGCAAGGAC AGCACCTACA GCCTCAGCAG CACCTGACG CTGAGCAAG CAGACTAGA GAACACAAA GTCTAGCCT CGGAAGTCA
 GTCTGCTCT GTGTTCTCTG TCGTGGATGT CCGAGTCTC GTGGACTCG GACTCGTTC GTCGATGCT CTTTGTGTT CAGATCGGA GCGTTCACTG

4801 CCATCAGGC CTGAGTCTG CCGTCAAAA GAGCTTCAAC AGGGAGGT GTTAAGTTC GATGGCCGC ATGGCCCAAC TTGTTTATG CAGCTTATAA
 GGTAGTCCC GACTCGAGC GGCAGTGTCT CTGCACTTC TCCCTCTCA CAATTCGAG CTACCCGCGG TACCGGTTG AACAAATAAC GTCSAATATTT

4901 TGGTTACAAA TAAAGCAATA GCATCAGAAA TTTTCAAAAT AAAGCAATTT TTTCACTGCA TTTAGTGT GTTTGTTG AAACAAATAA TGTATCTTAT
 ACCAATGTTT ATTTCTGTTT OGTAGTCTTT AAAGTGTTTA TTTCTGAAA AAAGTGAAGT AGATCAACA CCAACAGGT TTAGTAGT ACATAGAATA

5001 CATGCTGGA TCGGGAATTA ATTCCGCCA GCACRAGC CTGAATAAG TTTAACCCT CTGAAGAGG AACTTGTTA GGTACCCACT AGTAGCRAAG
 GTACAGACT AGCCCTAAT TRAGCCCGT CGTGGTACC GACTTATTC AATTTGGGA GACTTTCTCC TTGAACCAAT CCATGGCTGA TCATCTTCC

Figure 7-3

5101 TCGCCACGCA CAAGATCAAT ATTRACAATC AGTCATCTCT CTTTAGCAAT AAAAAGGTGA. AAAATFACAT TTTAAAATG ACACCATAGA CGATGATCA
AGCGTGGCT GTTCTAGTTA TAAATGTTAG TCAATAGAGA GAATGCTTA TTTTCCACT TTTAATGTA AAAATTTAC TGTGGTATCT GCTACATFAC

5201 AAATANCTA CTTGGAATA AATCTAGCA AGAAGTGA AGACTGTAC CCGAATAACT TACAATTTGT AAATGAGG TTAGTGAAGA TTTAAATGAA
TTTATTAGT GAACCTTAT TTAGATCCG TTCTTCCACT TCTGACAAAG GGTCTTTGA ATGTTTAA CA TTTACTCTCC AATFACCTCT AAATTTACTT

5301 TGAAGATCTA AATAACTTA TAAATTTGA GAGAAATTA TGAATGCTA AGTAAATGCA GAAACGGAGA GACATFACAT ATTCATGAAC TAAAGACTT
ACTTCTAGT TTAATTTGAT ATTTACACT CTCCTTAACT ACTTACAGAT TCAATTTACG CTTTGGCTCT CTTGATGATA TAAGTACTTG ATTTCTGAA

5401 AATATTGGA AGTATRCTT TCTTTTACA TAAATTTGPA GTCATATCT TCACCCCBAA AAAGCTGTTT GTTAACTTGT CAACCTCTT TCAAAATGTA
TTATAACACT TCCATTTGAA AGAAAGTGT ATTTAAACAT CAGTTATACA AGTGGGTTT TTTGACAAA CAATGGAACA GTTGGAGTAA AGTTTATCAT

5501 TATAGAAAGC CCAAGACAA TAACAAATAT ATCTTSTAG AACAAATGG GAAAGATCT TCCACTAAAT ATCAAGATTT AGAGCAAAGC ATGAGATGTS
ATATCTTTCC GGTTCCTGTT ATTTGTTTTA TAAGAACATC TTGTTTTACC CTTTCTTACA AGGTGATTTA TACTTCTAAA TCTCGTTTTG TACTCTACAC

5601 TGGGATAGA CAGTGGCT GATAAATAG AGTAGAGTC AGAAACAGAC CCAATGATAT ATGTAAGTGA CCAATGAAA AAATATGGCA TTTTACAATG
ACCCATCT GTCACTCGA CTATTTATC TCACTCGAG TCTTTGCTG GCTAACTATA TACATFACCT GGTACTTTTT TTTATACCGT AAATGTTTAC

5701 GAAATATGAT GATCTTTTC TTTTATGAA AARCGGGAA ATATATTTAT ATGTAATAAA TAAAGGGAA CCAATATGTC ATACCATACA CACAAAAAAA
CCTTTACTA CTAGAAAAG AAAAAATCT TTTCTCCCTT TATATAATA TACATTTTTT ATTTCCCTT GGGTATACAG TAAGGTATGT GTGTTTTTTT

5801 TTCCAGTGA TTATAAGTCT AAATGGGAA GCAAAACTT TAAATCTTTT AGAAATAAT ATAGAGCAT GCCATCATGA CTTCACTGTA GAGAAAAAT
AAGTCACTT AATATTCAGA TTTACTCTT CCGTTTTGAA ATTTAGAAA TCTTTTTATA TACTTCTGTA CCGTACTACT GAATCACAAT CTCTTTTTAA

5901 TCTTATGACT CAAAGTCTA ACCACAAGA AARGNTCTT AATTAGATG CATGAATAT AAGACTTAT TTTAAAATA AAAAACCATT AAAAAAGTC
AGATACTGA GTTTCAGAT TGGTGTCT TTTCTACAA TTAATCTAAC GTCTTATAA TTTGATAAA AAAATTTAAT TTTTGGTAA TTTCTTTCCAG

6001 AGCCATAGA ATGACAGAA ATATTTGCA CACCCAGTA AAGAAATG TAATATGAG AATTAFAAAA GAAGCTTAC AAATCAGTAA AAAATAAAA
TCCGGTACT TACTGTCTT TATAACGTT GTGGGTCAT TCTCTTAA ATATACCTC TAAATTTTTT CTTCAAGATG TTTACTCAT TTTTATTTG

6101 TAGACBAAA TTTGACAGA TGAAGAAA ACTTAATA ATCAATCAC ATGAAACT CAATCTCAGA AATCAGAAA CTATCATTCG ATATACACTA
ATCTGTTTTT AAATTTGCT ACTTTCTCT TGAATTTAT TAGTAATG TGCTTTGA GTTAGATCT TTAGTCTCTT GATAGTACG TATATGTGAT

6201 AATTAGAAA ATATTAAG GCTAAGTAC ATCTGGCA ATATTGATG TATATACCT TGAATGATG TGAATGAGAAC AGTACTTTAC CCAATGGCT
TTAATCTCT TATAATTTT CATTATCTG TAGACRCCG TATAACTACC ATATATGGA ACTATACTAC ACTACTCTTG CATGAAATG GGGTACCCG

6301 TCTCCCCA ACCCTTACC CAGTATAAT CATTGCAAT ATACTTTAAA AACATFACC CTATATCTAA CCACTACTCC TCAAACTCT CAAGTCTAC
AGGAGGGGT TGGGAATGG GTCATATTA GTACTCTTA TATGAATTT TTGSTRATG GATATAGAT GGTATGAGG AGTTTTGACA GTTCCAGTAG

6401 AAAAAFAGA AARGTCTGAG GAATCTCAA AACTAAGAG ACCCRAAGA GACATGAGAA TTATATGTA TGTGGCATTC TGAATGAGT CCCAGAACAG
TTTTTATCT TTTCAAGCT CTTGACACT TTGATTTCC TTGGTCTCC TTGGTCTCT CTGFACTCTT AATATACAT ACACCGTAAG ACTTACTCTA GGGTCTTCTC

6501 AAAAAAGCA GTAGTAAA AACTAATGAA ATATAATAA AGTTGACT TTAGTTTTT TTAARAAGA ATFCNAARA AATTTTTCT CATCGTAA GTGCGGTAAG
TTTTCTTGT CATCGATTT TTGATFACCT TATATTAAT TCAAACTGTA AATCNAARA AATTTTTCT CATCGTAA GTGCGGTAAG TCAATTTTCA

6601 ATTTTCTTG ACATTTAGT ACAATCTAT AATTAATAA TTTTAAATG TACTCTGAA CATTGCCAGA AACAGAGTA CAGCAGTAT CTGTGCTCTC
TAAAAAGAC TTGTAATCA TGTTCAGTA TTAATTTTA AAAAAATTTAC ATCAGACTT GTAACGCTCT TTTCTTCTAT GTCTGCGATA GACACGACAG

6701 GCCTAACTAT CCAATAGTGA TTGGTCTAAA ATGAGATACA TCAACCTCC TCCATGTTTT TTGTTTTCTT TTTAAATGAA AAATTTTAT TTTTAAAGAG
CGGATTTGATA GSTATGGACT AACAGATTT TACTCTATGT AGTTGGCAGG AGTACAAAA AACAAAGRA AAATTTACTT TTTGAAATAA AAAATTTCTC

6801 AGTTTCAGT TCAATAGCAA ATTAGAGGA AGGTACATTC AAGCTGAGGA AGTTTCTCT TATTCCTAGT TTAGTGAGG ATTCATCAT GAATGGGCT

Figure 7-4

TCAAAGTCCA AGTATCGGTT TAACTCTCCT TCCATGTATG TCGACTCCT TCAAAAGGAG ATAGGATCA AATGACTCTC TAACTGTAGTA CTTACCCACA
 6901 TAAATTTTGT CAAATGCTTT TCTGTGTCT ATCAATATGA CAGTGTGAT TCTCTCTTTA ACCTGTTGAT GGGACAAAT ACCTGTTATG ATTTTCRAAC
 ATTTAAACA GTTTACGAAA AAGACACAGA TAGTTATACT GGTACACTAA AAGAAGAAAT TGGCAACTA CCTGTGTTAA TGCAAATTAAC TAAAAAGTTTG
 7001 GTTGAACAC CCTTACATAT CTGGATAAA TCTACTTGG TTGGTGTGA TATTTTGA TACATCTTGG GATCTTTTGT GCTAATATTT TGTTCGAAAAT
 CAACTTGGT GGAATGTATA GACTTATTT AAGATGAACC AACACCACAT ATAAAABACT ATGTAAGAAC CTAAGAAAAA CGATTATRAA ACAACTTTTA
 7101 GTTTGTATCT TTGTTATGA GAGATATGG TCTGTGTTT TCTTTCTTG TAAATCAAT TCTAGTTCC GGTATPAAGG TAAATGTGGC CTAGTTGAAT
 CAAACATAGA AACAAGTACT CTCTATAACC AGACAACAAA AGRARAAGAC ATTACAGTAA MAGATCAAGG CCATAATTTCC ATTTACGACCG GATCAACTTA
 7201 GATTTAGAA GTATTCCTC TGTCTGTG TCTGAGTA CCGCGCCGC CCGTGTGTTT ACAACGTCTG TACTGGGAAA ACCCTGGCGT TACCGAACTT
 CTAATCTCTT CATPAGGAG AGRAGACAG AAGACTCCAT GCGCGCGGG GGCACRARA TGTTCAGCA CTGACCCCTT TGGGACCGCA ATGGGTTGRA
 7301 ARTGCCCTG CAGCACATCC CCCCCTGCC AGCTGGCTA ATAGGAA GAACCGCAC GATCGCCCTT CCCAACAGTT GGCAGCGCTG AATGGCGGAAT
 TTAGCGGAAC GTGTTGTAGG GGAARAGGG TCGACCGCAT TATCGTTCT CCGGGCGTGG CTAGCGGAAA GGTGTGTCAA CCGGTGGGAC TTACCGCTTA
 7401 GCGCCCTGAT GCGTATTTT CTCTTACG ATCTGTGGG TATTTACAC GGCATACGTC AAGCAACCA TATGACGCGC CCTGTAGCGG CGCATTAAGC
 CCGCGACTA CGCCATAAA GAGGATGCG TAGCACGCC ATAAATGTC GGTATGTCAG TTTGTTGT ATCATGCGCG GGACATCGCC GCGTATTTCC
 7501 GCGCGGGTG TGTGGTTAC GCGCAGCGTG ACGGTACAC TTGCCAGGC CCTAGCGCC GCTCCTTTGG CTTCTTTCC CTTCTTTCC GGCACGTTGG
 GCGCGCCAC ACCACCAATG CCGCTCGCAC TGGCATGTC AACGATGTC GATCAAGAA ATGCGTGGG GATGAAAGG GAAAGAAAG CCGTGCRAAGC
 7601 CCGCTTTCC CCGTCAAGCT CTAATCGGG GGTCCCTTT AGGTTCCGA TTAGTCTTT TACGGCACCT CGACCCCAA AAATTTGAT TGGGTGATGG
 GCGGARAAG GCGATTCGA GATTTAGCC CCGAGGAAA TCCCRAGCT AAATCAGAA ATGCGTGGG GCTGGGTTT TTTGAACATA ACCCACTACC
 7701 TTCAGTAGT GCGCATCGC CTGATAGC GGTTTTCCG CTTTGAAGT TGGATCCAC GTTCTTTAAT AGTGGACTCT TGTTCGAAA TGGAAACAAC
 AAGTGATCA CCGGTAGCG GACTATCTG CCRAAAGCG GAAACTGCA ACCTCAGTGC CARGAATTA TCACCTGAGA ACAAGTTTG ACCTTTGTTGT
 7801 CTCACCCCA TCTGGGCTA TTTTCTGAT TTATAAGGA TTTTCCGAT TCGGCTAT TGGTAAAAA ATGAGTGTAT TTAACAADAA TTTAAACGCGA
 GAGTTGGAT AGACCCGAT AGRANAATA AMATTTCTT AAACGGGTA AAGCGGATA ACCAATTTT TACTCGACTA AATTTGTTTT AAATTTGCGCT
 7901 ATTTTACAA AATATRACG TTTACATTT TATGTTGAC TCTCAGTACA ATCTGCTCG ATGCCGATA GTTAAGCCAG CCCCAGACCC GGCACAACALC
 TAAATTTGTT TTATATTTGC AANTGTTAAA ATCCACGTC AGAGTCTGT TAGACGAGAC TACGGGTTA CAATTCGGTC GGGGCTGTGG CCGTGTGGOI
 8001 CCGTACGCG CCGTACGCG CTTGCTGCT CCGGATCC GGTACAGAC AAGTGTGAC CGTCTCCGG AGCTGCATGT GTCAGAGTT TACACCGTCAV
 GCGACTGCG GCGACTGCC GAACAGAGA GGGCGTAGS GAAATGTCG TCGACACTG SCAGAGGCC TCGRGTACA CAGTCTCAA AAGTGGCAGT
 8101 TCACCGAAC GCGCGRAGA GAAAGGCC TCGTATAGC CCTATTTTA TAGTATATG TCATGATAAT AATGTTTCT TAGAGTCTAG GTGGCCTTTT
 AGTGGCTTTG CCGCTCTCT GCTTTCCGG AGCACTATG GATTAARAAT ATCCAATTAC AGTACTATTA TACCARAAGA ATCTGCAGTC CACCGTGA
 8201 TCGGGAAAT GTGGCGGAA CCCCATTATG TTTATTTTC TAAATACAT CAAATATGTA TCCGCTCATG AGACATAAC COTGATAAT GCTTCAATAA
 AGCCCTTTA CACGGCCTT GGGATRAAC AATPAAAAG ATTTATPAA GTTTATACAT AGGCGAGTAC TCTGTTATTT GACTATTTA CGAAGTTAT
 8301 TATTAAGAA GAAAGATAT GATATTCAA CATTTCCGTC GCGCCTTAT TCCCCTTTT GCGGATTTT GCTTCTCTAC AAACGAGTG GGTCTTTGCG
 ATAACTTTT CCTTCTATA CTCAATAGTT GTRAGGAC ACAGGATA AGGAAAAA GCGCTTAAA CCGAAGGACA AAACGAGTG GGTCTTTGCG
 8401 TGGTGAAGT AARAGTCT GAAGTACG TGGTGCAC AGTGGTATC ATCTCAGAC CCGTARAC GGTGAGATC CTGAGAGTT TCGCCCTGGA
 ACCACTTTCA TTTTCTACGA CTTCTAGTCA ACCACGTC TCACCCRAATG TAGGTTGTC GCCATTTAG GAACTCTAA AAGCGGGCT
 8501 AGRACGTTTT CCAATGATGA GCACTTTTAA AGTCTGTGTA TGTGGCGGG TATATCCCG TATTACGCC GGGCAGGAC AACTCGGTC CUCCATACAC
 TCTTGAACA GGTACTACT CGTGAATAAT TCAAGACGAT ACACCGGCC ATAAATAGCG ATAACTGCG CCGTCTCG TTGACCCAG CCGTATGTTG

Figure 7-5

8601 TATTCTCAGA ATGACTTGGT TGAGTACTCA CCAGTCACAG AAAAGCATCT TAGGATGCC ATGACAGTAA GAGAAATTATG CAGTCTGCC ATAACCATGA
ATAAGAGTCT TACTGRACCA ACTCTGAGT GGTCAAGTGC TTTTCGTAGA ATGCTTACCG TACTGTCAAT CTCCTAATAC GTCACGACGG TATTGTACT

8701 GTGATPAACAC TCGCGCCAAC TTACTTCTGA CACGATCGG AGACCCGAG GAGCTAACCG CTTTTTTCGA CARGATGGGG GATCATGTAA CTCGCCCTTGA
CACTATTGTG ACGCCGGTTG AATGAAAGACT GTTGCTAGCC TCTGGCTTC CTGAAATGGG GAANAACGGT GTTGTACCCC CTGTACACTT GAGCGGAAC

8801 TCGTTGGGAA CCGGAGCTGA ATGAGCCAT ACCAAGCAC GAGGTGACA CCAGATGCC TGTAGCAATG GCACAAAGT TCGCAAACT ATTAAGTGGC
AGCAACCCCTT GGCCTCGACT TACTTCGGTA TGGTTGCTG CTGCACTGT GGTCTAGG ACATCGTTAC CGTGTTCGA ACGGCTTGA TAATTGACCG

8901 GAACTACTTA CTCCTAGCTTC CCGGCAACA TTAATAGACT GATPAGAGG GATPAAAGTT GCAGACAC CTTCTGGCTC GGCCTTCCG GGTGGTGGT
CTTGATGAT GAGATCGAG GSCCGTGT TAAATNCTGA CCTACTCCG CCIATTTCAA CTTCTGCTG AAGACCGAG CCGGNAAGC CGACCCAGCA

9001 TTAATGCTGA TAAATCTGA GCGGTGAGC GTGGTCTCG CGTATCAAT GCAGACTGG GGCAGATGG TAAGCCCTCC CBTATCGTAG TTAATCTACAC
AATAAGACT ATTAGACCT CGGCCACTCG CACCCAGAC GCCATAGTAA CTTGCTGACC CCGTCTTACC ATTCGGGAG GCATAGCATC AATAGATGTG

9101 GACGGGAGT CAGGCAACTA TGGATGAACG AATAGACAG ATCGCTGAGA TAGTGCCTC ACTGATTAAG CATTGGTAACT TGTGAGACCA AGTTTACTCA
CTGCCCTCA GTCCGTYGAT ACTACTTGC TTTATCTGTC TAGGACTCT ATCCACGGAG TGACTAATC GTAACCATG ACAGTCTGT TCAAATGAST

9201 TATATACTTT AGATTGATTT AAACTTCAT TTTAATTTA ARAAGTCTA GGTGAAGTC CTTTTGATA ATCTATGAC CAAAATCCCT TAACGTGAGT
ATATATGAA TCTRACTPAA TTTTGAAGTA AAAATPAAI TTTCTTAGT CCACTTCTAG GAAAACATAT TAGAGTACTG GTTTTAGGGA ATTGCACTCA

9301 TTTCTGTTCA CTGAGGCTCA GACCCGCTAG AAAAGTCAA AGGATCTCT TGAGATCTT TTTTCTGGC CBTAACTGC TGTGTGAAA CAAAAAAC
AAAGCAAGT GACTCGCAT CTGGGCATC TTTCTAGTT TCTTAGA ACTTAGA AAAAGACGC CANTAGACC ACGAACGTT GTTTTTTGG

9401 ACCGTTACCA GCGGTGGTT GTTTGCCGA TCAAGACTA CBACTCTT TCCGAAGT AACTGGCTC AGCAGAGCC AGATACCANA TACTGTCTT
TGGCATGT CCGCACCAA CAACGGCT AGTCTCGAT GPTTAGANA AAGCTTCCA TTGACCCAG TCGTCTGCG TCTATGGTT ATGACAAGAA

9501 CTAGTCTAG CBTAGTTAG CCACACTC AAGACTCTG TAGCAGCC TACTACTC GCTCTGCTAA TCTTGTACC AGTGGCTCT GCCAGTGGC
GATCAGATG GCATCAATC GGTGTGAG TTCTTGTAGC ATCGTGGCG AUCTATGAG CGAGACGAT AGGACATGG TCACCCGACA CGGTACCCG

9601 TAAATCTGT TCTTCCGGG TTGGCTCA GACATAGTT ACGGATAAG GCGCAGCGT CCGCTGAAC GGGGGTTCG TGCACACAGC CCAGCTTGA
TATTACAC AGAATGGCC ACCCTGAT CTGCTATCAA TGGCTATTC CCGCTGCCA CCGCAGCTT CCGCCCAAGC ACGTGTGTG GGTCAACCT

9701 GCGAAGACC TACCCGAC TGAGATACCT ACAGCTGAG CTATGAAA GCGCCAGCT TCCCGAAGG ABRAGCGG ACAGGTATCC GGTAAAGCC
CGCTTCTGS ATGTGGCTG ACTCTANGA TGTGCACTC GATCTCTT CCGGCTGCA AAGGCTTCC TCTTCCGC TGTCCATAGG CCAATCGCC

9801 AGGTCCGAA CAGGAGCG CAGGAGGAG CTTCAGGGG GAACGCTG GATCTTTAT AGTCTGTG CAGGTTGCCA CCTCTGACTT GAGGTCTGAT
TCCGAGCCTT GTCTCTGCG GTGCTCCCTC GAAGTCCCC CTITGGGAC CATAGAAATA TCAGGACAGC CCAAGCGGT GAGACTGAA CTCGAGCTA

9901 TTTTGTGATG CTCGTCAGG GGGGGAGCC TATGAAA AA GCGCCGTAAC GCGGCTTTT TAGGTTCTT GGCCTTTGC TGGCTTTTG CTCACATGTT
AAACACTAC GAGCAGTCC CCGCCTCG ATACTTTTT GGGTGTGTT CCGCCGAAA ATGCCAAGGA CCGGAAAACG ACCGGAAC GAGGTACAA

10001 CTTTCTGCG TTATCCCTG ATTTGTGGA TAACTGAT TACCGTAT ACCGCTTTG AGTAGCTGA TACCGTCCG CCGAGCCGAA CAGCCGAGCG CAGGAGTCA
GAAAGACCG AATAGGGAC TAAGACACT ATTTGATRA TGGCGAATC TCACCTGACT ATGGCGAGC GGTCTGCTG GCTGGCTCG CTCCTCAGT

10101 GTGAGGAGG AAGCGAAGA GCGCGGGG AAGTGGCCA CCGCAAGAT CAATATTAAC AATCACTCAT CTCTTTTAG CAATAAAG GTCAAAAATTT
CACTGCTCC TTCGCTTCT CCGGCCCCG TTCCAGGCTT GGTGTCTTA GTTATPATTG TTAGTCAGTA GAGAGAATC GTTATTTTC CACTTTTTAA

10201 ACATTTTAA AATGACCA TAGAGATGT ATGAAATAA TCTACTTGA AATAATCA GCAARAAG TGCAGACTG TTAACCCAGAA AACTTACAAA
TGTAAAATTT TTAGTGTGT APTCTTCA TACTTTTAT AGATGAACCT TTTTATAGAT CCGTTTCTC ACGTCTGAC AATGGTCTT TGAATGTTT

10301 TGTAAATGA GAGTTAGTG AAGATTTAA TGAATGAGA TCTAATPAA CBTAAAT GTGAGAGAA TAAATGATG TCTAAGTTAA TGCAGAAACG
AACATTTACT CTCATAC TCTAATTT ACTTACTTCT AGATTTATTT GAATTTTAA CACTCTCTT AATTACTTAC AGATTTCAAT ACCTCTTTCG

Figure 7-6

10401 GAGAGACATA CTATATTCAT GAACTAABAG ACTAATATT GTGAGGAT ACTTCTTTT CACATAAATT TGTAGTCAAT ATGTTCCACC CAAAAAGCT
 CICTCTGTAT GATATRAAGTA CTTGANTTTC TGAATATPAA CACTTCCATA TGAAGAAAAA GTGATTTTAA ACATCAGTTA TACAAGTGGG GTTTTTTCGA

10501 GTTTGTTAAC TTCTCAACCT CATTTCRAAA TGTATATAGA AAGCCCAAG ACAATAACAA AATATTTCTT GTAGAACAAA ATGGGAAGA ATGTTCCACT
 CAACAATG AACAGTTGGA GTAAAGTTTT ACATATATCT TTGGGTTTC TGTATTTGTT TTTATAAGAA CAICTTGT TTCCCTTTCT TACAAGGTGA

10601 AATATCAAG ATTTAGACA AAGCATGAGA TGTGTGGGA TACACAGTGA GCCTGATPAA ATAGATPAGA CACTGAAAC AGACCCATG ATATHTGTAA
 TTTATAGTTC TAAATCTGT TTCTACTCT ACACCCCT ATCTGTCACT CCGACTATTT TATCTCACT CGAGTCTTIG TCTGGTARAC TATATACATT

10701 GTGACCTAG AAAAAAATHT GGCATTTTAC AATGGGAAA TGAATCTTT TTCTTTTTT AGAARACAG GGAATATPAT TTTATGTPAA AAAATAAAG
 CACTGATAC TTTTTTATA CCGTAAAATG TTACCCTTT ACTACTAGAA RAAGAAAAA TCITTTTCTC CCTTATATA AATATACATT TTTTATTTTC

10801 GGAACCCATA TGTCAVACCA TACACACRAA AAAAAATCCAG TGAATATPAA GTCTAATGG AGRAGGAAA ACTTTAATC TTTTAAAAA TAATATPAGAA
 CCTGGGTAT ACAGTATGTT ATCTGTCTTT TTTTRAGTCT ACTAATATTT CAGATTTACC TCITCCGTTT TGAATTTTAG AATATCTTTT ATATATCTT

10901 GCATGCCATC ATGACTTCAG TGTAGAGAAA AATTTCTTAT GACTCAAAGT CCTAACACA AAGAAAGAT TGTATATPAG ATTCATGAA TATTAAGACT
 CSTRCCGTAG TACTGAATC ACATCTCTTT TTAAGAAATA CTGAGTTTCA GGATTTGTTT TCTTTTCTA ACATTAATC TAACTACTT AATATCTGA

11001 TATTTTAA ATTAABAAAC CATTAAGAA AGTCAGGCCA TACATGACA GAAATATTT GCAACCCCT AGTAAAGAGA ATTTAATAT GCAGATATTA
 ATAAAAATTT TAAATTTTGG GTAATTTCTT TCAGTCCGGT ATCTACTGT CTTTTATAA CGTTGGGG TCATTTCTCT TACATATATA CGTCTAATAT

11101 AABRAGTCT TTACAAATCA GTAAAAATA AACTAGACA AAAAAATGAA CAGATGAAAG ARAACTCTA AATATCAAT ACATGAGAGA AACTCAATCT
 TTTTCTTCAG AATGTTTATG CATTTTTAT TTTGNTCTGT TTTTAACTT GTCTACTTTC TCTTTGAT TTAATGTPAA TGTACTCT TTGAGTAGA

11201 CAGAAATCAG ABACTATCA TTCCATATAC ACTAATATG AAAAAATTA AAGGCTAAG TAACATCTGT GCAATATTT ATGATATATA ACCTTGATAT
 GCTTTTATG TCTTGATAGT AACGTATATG TGAATTAATC TCTTATAAT TTTCCGATTC ATTTAGACA CCGTATAAC TACCATAT TGGAACTATA

11301 GATGTGATGA GAACACTT TTACCCCATG GGTCTCTCC CCAACCCCT ACCCCAGTAT AATATATGAC AATATATCTT TAAARCCAT TACCCTATAT
 CTACACTACT CTGTCTATGA ATGGGGTAC CCGAAGGAG GGTTTGGAA TGGGTCTATA TTTAGTACTG TTTATATGAA ATTTTGGTA ATGGATATA

11401 CTARCCAGTA CTCTCRAAA CTGTCAAGGT CATCAAAAT AAAAAAGTC TGAGAACTG TCAAACTPAA GAGAAACCA AGGAGCATG AGAATATAT
 GATTTGTCAT GAGGATTTT GACAGTTCCA GTAGTTTTTA TTCTTTTTCAG ACTCTTGC ACITTTGATTT CTCCCTGGGT TCCTCTGTAC TCTTAAATATA

11501 GAAATGTGC ATTTGAAAG AGATCCAGA ACAGAAAG ACAGTACTT ARAACTATA TGAATATPAA ATAAAGTTG AACTTTACTT TTTTTTAAAT
 CAATACCCG TAAGACTTAC TTAGGGTCT TGTCTTTTTC TTGCAFCGA TTTTITGATTT ACTTATATTT TATTTCAAC TTGAAATCAA AAAAAATTTT

11601 AAGGTAGCA TTAACAGGC AAGTCAATTT TCATATTTT CTGACATTT AATGTCATGTA TTTTAAAAA ATATTTTAA AATGTCATTT GGAACATTC
 TTCTCATCT AATTTGCGG TTTCACTAAA AGTATAAAA GAACTTTGTA TCTGTGTCA GATATTAAT TTTAAAAAAT TTACATCAGA CCTTGTACCG

11701 CAGAACAGA AGTACAGCAG CTATCTGTGC TGTCCCTPAA CTATCCATG CTGATGGTC TAAATGAGA TACATCAAG CTCCTCCATG TTTTTGTTTT
 GTCTTTGCT TCACTGTCTC GATAGACAG ACAGCGAAT GATAGTATC GACTAACCCAG ATTTTACTCT ATGTAGTTGC GAGAGGTAC AAAAAACAAA

11801 TCTTTTTAA TGAARAACTT TATTTTTAA GAGGATTC AGTTCCATAG CAAPTTGAG AGGAGTAC ATTCRAAGT AGGAAGTTT CCTCTATTC
 AGAAAAATTT ACTTTTTGAA ATAAAAATTT CTCTCAAG TCCAAATATC GTTTTAACTC TCCTTCCATG TRAGTTGAC TCCTTCAAA GGAGATAAG

11901 TAGTTTACTG AGRATTTGA TCAATGATGG GTGTTAATTT TTGTCATATG CTTTTTCTGT GTCATCAAT ATGACCATGT GATTTTCTC TTTTAACTCT
 ATCAATGAC TCTTTRACGT AGRCTTACC CACAATTTA RACGTTTAC GAARAGACA CAGATAGTTA TACTGGTACA CTAAAGAG AAATTTGACA

12001 TGATGGACA AATTAGTTA ATTTGTTTTT AAACGTTGAA CCACCCCTAC ATATCTGAA TAAATTTCTAC TTGGTGTGG TGTATATTT TTGATACATTT
 ACTACCCTGT TTAATGCAAT TAACTAAAAG TTGCAACTT GGTGGGAATG TAPAGACTT ATTTAAGATG ACCAACACC ACATATPAAA AACTATGTA

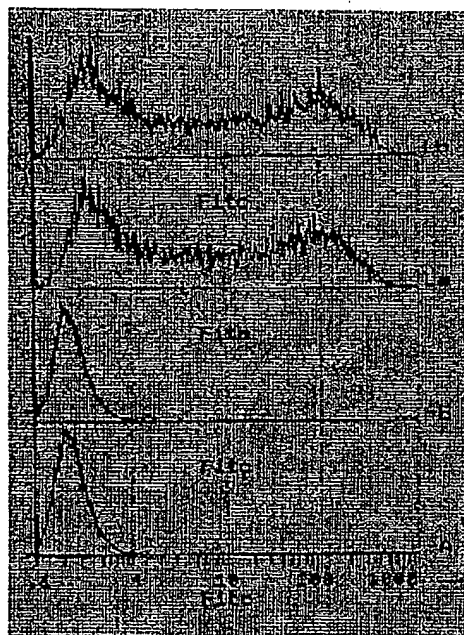
12101 CTGGATCT TTTTGTAT ATTTTGTGA AATGTTTTGT AICTTTGTTT ATGAGATA TTGGTCTGT. GTTTTCTTTT CTGTATATGT CATTTTCTAG

Figure 7-7

GAACCTAAGA AARCGAATTA TAAACAACACT TTTACAAACA TAGAACAAAG TACTCTCTAT AACAGACAA CAARAGAAA GRACATTACA GTAAAGATC
 12201 TTCGGTATT AAGGTAAAGC TGGCCTAGTT GAATGATTTA GGAAGTATTC CCTCTGCTTC TGTCTTCTGA AGGGGAAGAG CGCCCAATAC GCAAACCCGCC
 AAGGCCATAA TTCCATTACG ACCGGATCAA CTTACTAAT CTTTCATAAG GGAGACGAAG ACAGAAGACT TCCCTTCTC GCGGGTTATG CGTTTGGGGG
 12301 TCTCCCGCG CGTTGGCCGA TTCATTAATG CAGCTGGCAC GACAGTTTC CCGACTGGAA AGGGGCACT GAGGGCAACG CAATTAATGT GAGTTAGCTC
 AAGGGGGGC GCRACCGGCT AAGTAATFAC GTCGACCGTG CTGTCCAAAG GGTGACCTT TCCCCCGTCA CTCGCGTTGC GTTAATTACA CTCAAATCGAG
 12401 ACTCATTAGG CACCCAGGC TTTACACTTT ATGCTTCCGG CTCCTATGTT GTGTGGAAAT GTGAGGGAT AACAATTTCA CACAGGAAC AGCTATGACA
 TGAGTAATCC GTGGGTCCG AARTGTGAAA TACGANGGC GAGCATACAA CACACCTTAA CACTCGCTA TTGTTAAGT GTGTCTTTG TCGATACTGT
 12501 TGATTACGAA TTAA
 ACTAATGCTT AATT

>length: 12514

Figure 7-8



<u>% GFP</u>	<u>% Viability</u> (PI Staining)
70.9	59.2
68.7	60.9
1.6	69.7
1.2	94

Figure 8. FACS analysis of transiently transfected CHO cells with a GFP plasmid in 250 ml spinner transfection.

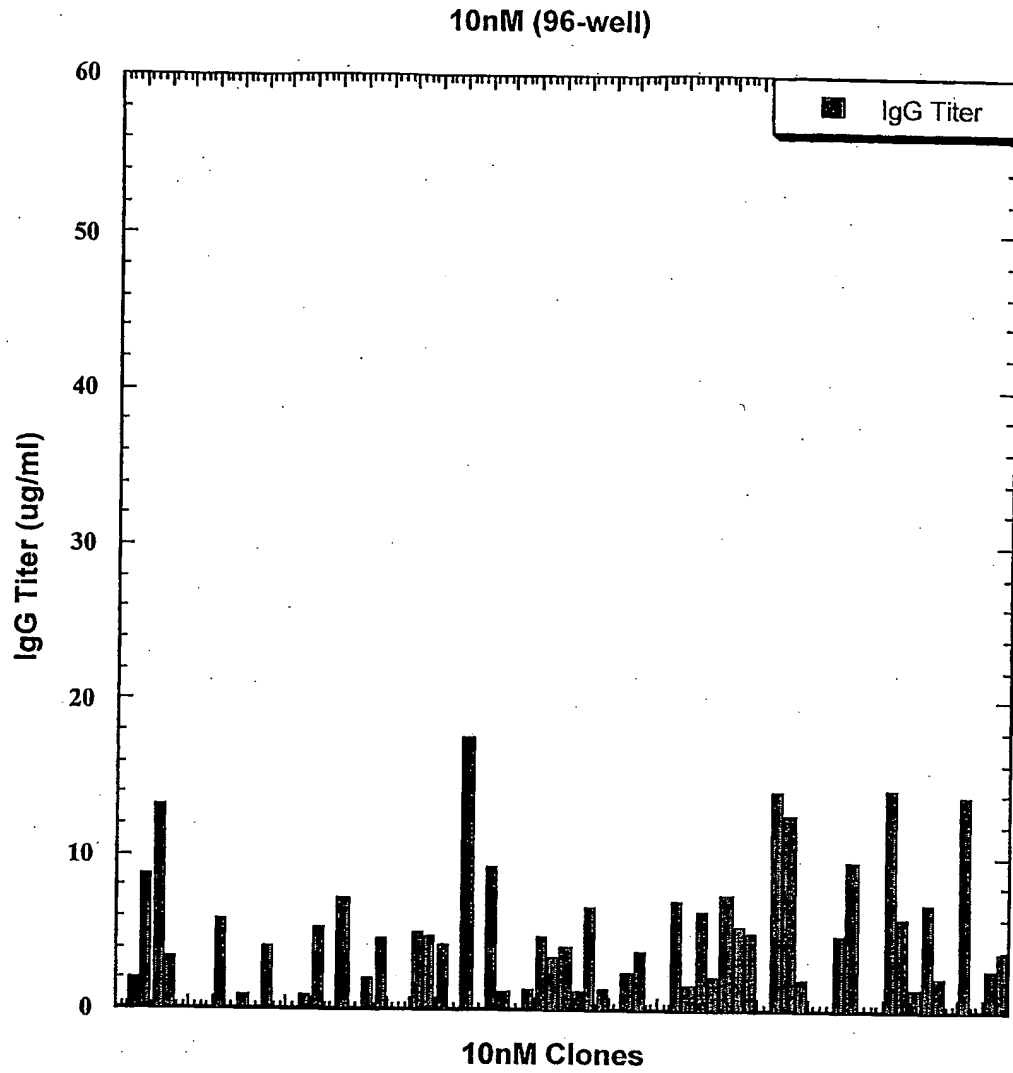


Figure 9. Expression level of clones from traditional 10 nM MTX selection.

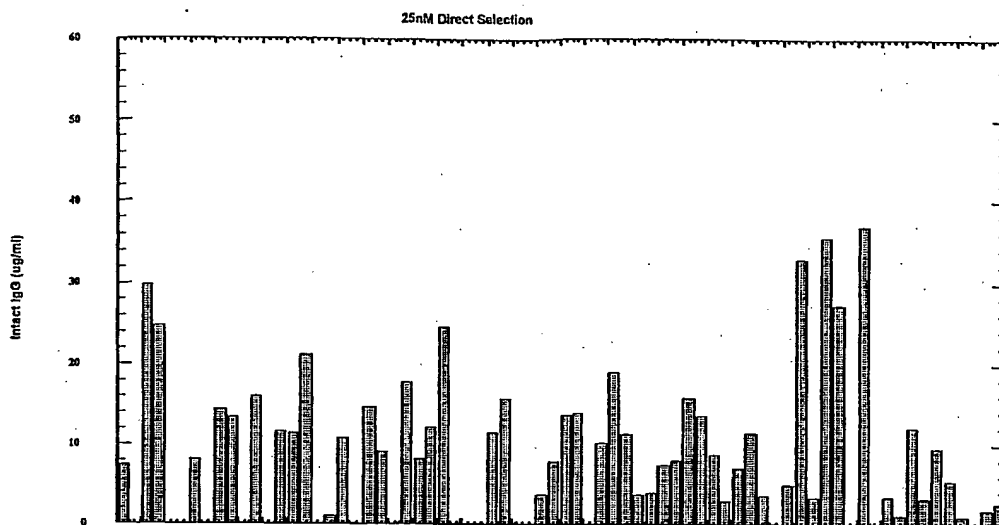


Figure 10-1

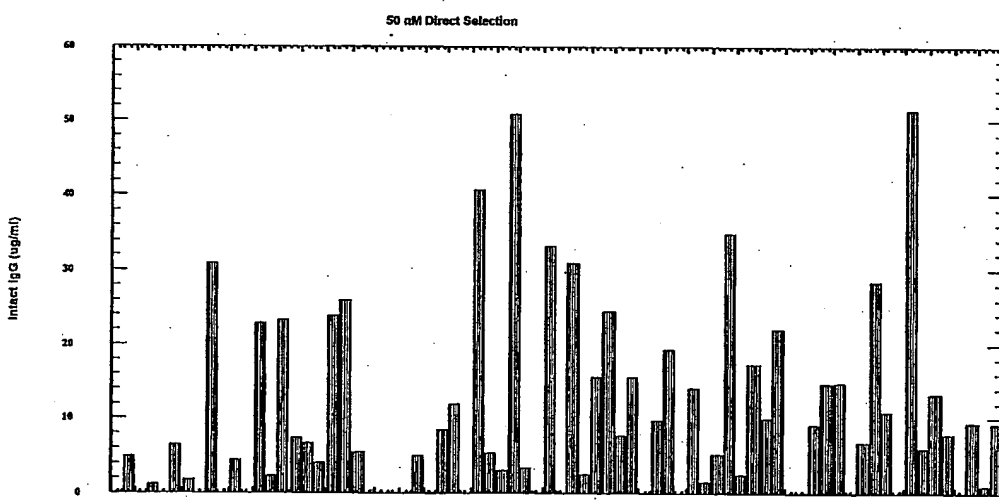


Figure 10-2

Figures 10.1 and 10.2. Expression level of clones from 25 and 50 nM MTX direct selections of SV40-based constructs derived from spinner transfection, respectively.

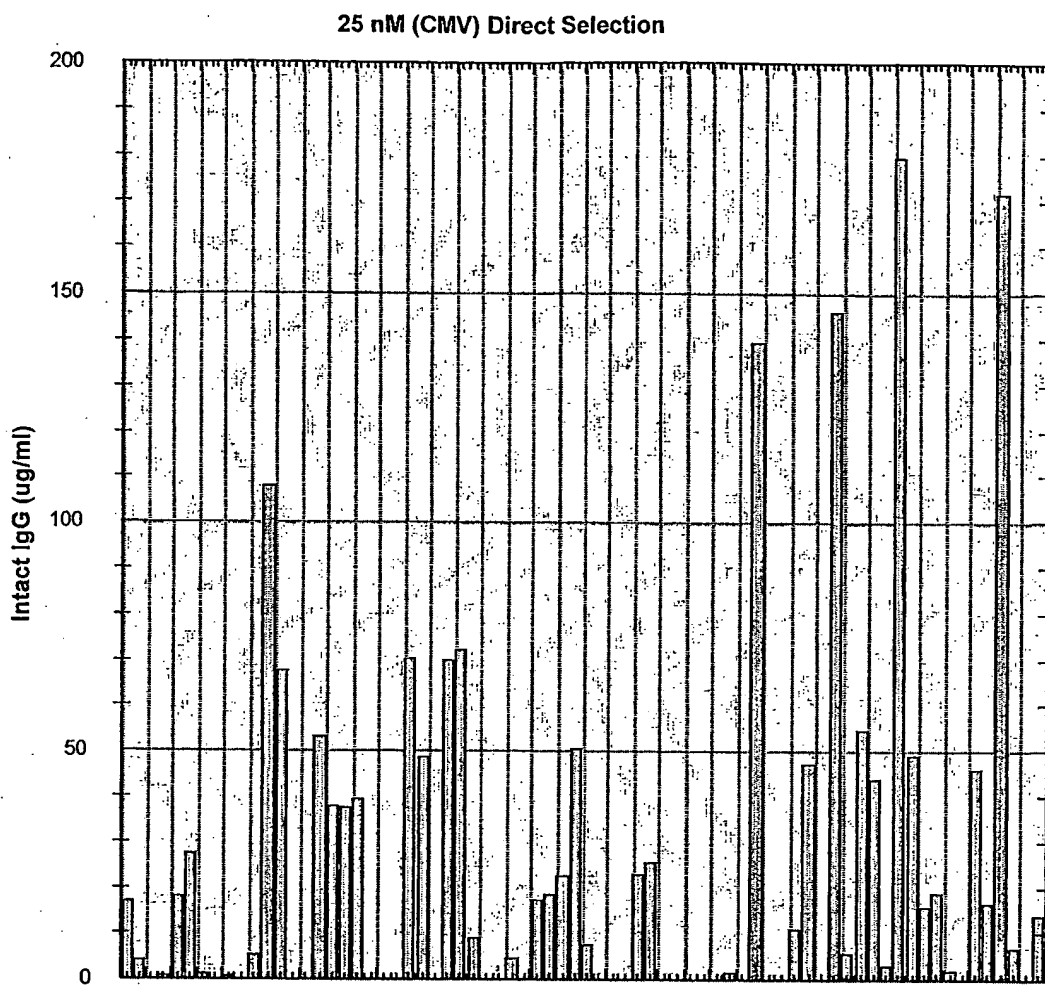


Figure 11. Expression level of clones from 25 nM MTX direct selection of CMV construct derived from spinner transfection.

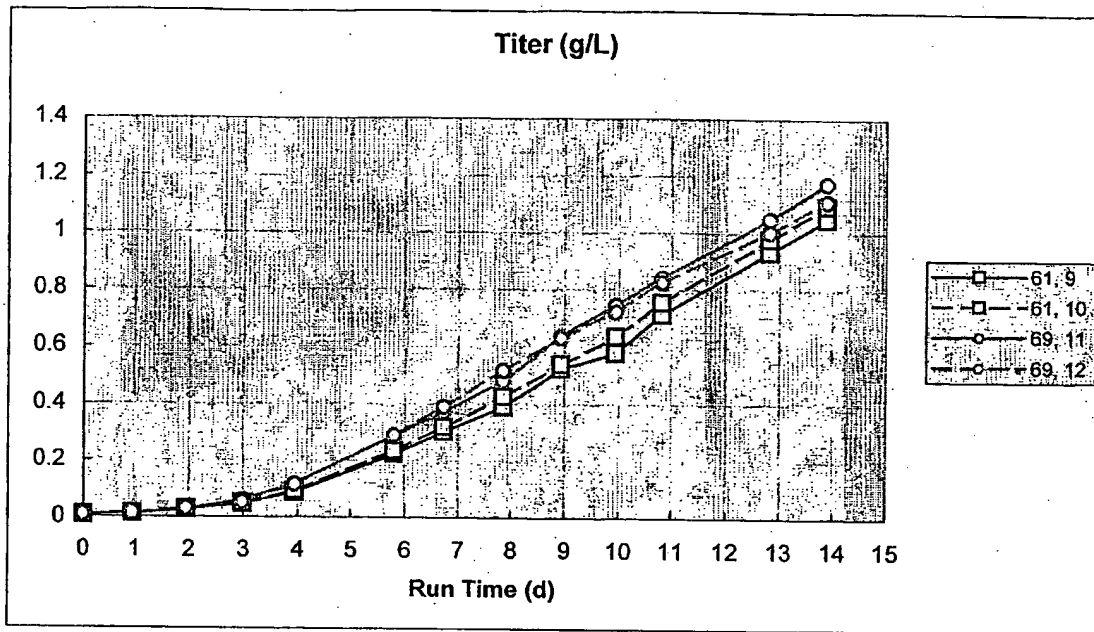


Figure 12. Titer Evaluation in Miniferm.

Figure 13. Plasmid pCMV.IPD.Heterologous Polypeptide

5 <400>
60 TTCGAGCTCG CCCGACATTG ATTATTGACT AGAGTCGATC ACCGGTAGTA ATCAATTACG
120 GGGTCATTAG TTCATAGCCC ATATATGGAG TTCCGGGTTA CATAACTTAC GGTAAATGGC
180 CCGCCTGGCT GACCGCCCAA CGACCCCGC CCATTGACGT CAATAATGAC GTATGTTCCC
240 ATAGTAACGC CAATAGGGAC TTTCATTGA CGTCAATGGG TGGAGTATT ACGGTAAACT
300 GCCCACCTGG CAGTACATCA AGTGATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT
360 GACGGTAAAT GCGCCGCCCTG GCATTATGCC CAGTACATGA CCTATATGGA CTTTCCTACT
420 TGGCAGTACA TCTACGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT TTGGCAGTAC
480 ATCAATGGGC GTGGATAGCG GTTGACTCA CGGGGATTTC CAAGTCTCCA CCCCATTTGAC
540 GTCAATGGGA GTTGTGTTTG GCACCAAAAT CAACGGGACT TTCCAAAATG TCGTAACAAC
600 TCCGCCCCAT TGACGCAAT GGGCGGTAGG CBTGTACGGT GGGAGGTCTA TATAAGCAGA
660 GCTCGTTTAG TGAACCGTCA GATCGCCTGG AGACGCCATC CACGCTGTTT TGACCTGGGC
720 CCGGCCGAGG CCGCCTCGGC CTCTGAGCTA TTCCAGAAGT AGTGAGGAGG CTTTTTTGGA
780 GGCCTAGGCT TTTGCAAAA GCTAGCTTAT CCGGCCGGGA ACGGTGCATT GGAACGGCGGA
840 TTCCCGGTGC CAAGAGTGAC GTPAAGTACCG CCTATAGAGC GACTAGTCCA CCAFGACCGA
900 GTACAAGCCC ACGGTGCGCC TCGCCACCCG CGACGACGTC CCGGGGGCCG TAGGCACCCCT

Figure 13.1

960 CGCCGGCCGG TTGCGCGACT ACCCGGCAC GCGCCACACC GTAGACCCGG ACCGCCACAT
1020 CGAGCGGGTC ACCGAGCTGC AAGAACCTTT CCTCACGGCG GTGCGGGCTCG ACATCGGGAA
1080 GGTGTGGGTC GGGACGAGG GCGCGGGCGT GCGGGTCTGG ACCACGCCGG AGAGCGTCGA
1140 AGCGGGGGG GTGTTGCGCG AGATCGGCC GCGCATGGCC GAGTTGAGCG GTTCCCGGGCT
1200 GGC CGCCGAG CAACAGATGG AAGCCCTCTT GCGCCGGCAC CGGCCAAGG AGCCCGGGTG
1260 GTTCTGGCC ACCGTGCGG TCTCGCCGGA CCACCAGGCG AAGGGTCTGG GCAGCGCCCT
1320 CGTGTCCCC GGAGTGGAGG CGGCCGAGCG CGCCGGGGTG CCCGCCITCC TGGAGACCTC
1380 CGCGCCCGC AACCTCCCT TCTACGAGG GCTCGGGTTC ACCGTACCG CCGACGTCGA
1440 GGTGCCCCGAA GGACCCGGCA CCTGGTGCAT GACCCGCAAG CCCGGTGCCA ACATGGITCG
1500 ACCAATGAAC TGCAATGTCG CCGTGTCCCA AAATAIGGG ATTGGCAAGA ACGGAGACCT
1560 ACCCTGGCCT CCGCTCAGGA ACGCGTTCAA GTACTTCCAA AGAATGACCA CAACCTCTTC
1620 AGTGAAGGT AAACAGAATC TGGTGAATAT GGGTAGGAAA ACCTGGTTC CCATTCTTGA
1680 GAAGRAATCGA CCTTTAAAGG ACAGAAATTA TATAGTTCTC AGTAGAGAC TCAAGAAACC
1740 ACCACGAGGA GCTCAATTC TTGCCAAAAG TTTGGATGAT GCCTTAAGAC TTATTGAACA
1800 ACCGGAATG GCAAGTAAAG TAGACATGGT TTGGATAGTC GGAGGCAGTT CTGTTTACCA
1860 GGAAGCCATG AATCAACCAG GCCACCTCAG ACTCTTTGTG ACAAGGATCA TGCAGGAAT
1920 TGAAGTAC ACGTTTTTCC CAGAAATTA TTTGGGAAA TATAACCTC TCCCAGATA
1980 CCCAGCGTC CTCTCTGAGG TCCAGGAGGA AAAAGGCATC AAGTATAAGT TTGAAGTCTA

Figure 13.2

2040 CGAGAAGAAA GACTAACGTT AACGTGTCCTCCTAAAGC TATGCATTTT TATAAGACCA
 2100 TGAGACTTTT GCTGGCTTTA GATCCCTTG GCTTCGTTAG AACGCAGCTA CAATTAATAC
 2160 ATAACCTTAT GTATCATACA CATACGATTT AGGTGACACT ATAGAATAAC ATCCACTTTG
 2220 CCTTCTCTC CACAGGTGTC CACTCCAGG TCCRACTGCA CCTCGGTTCT ATCGATTGAA
 TTCCACC --Insert Sequence of Interest--
 CGA TGGCCGCCAT GGCCCAACTT GTTTATTGCA GCTTATAATG
 GTTACAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT
 CTAGTTGGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGCTGGATC GSGAATTAAT
 TCGGGCAGC ACCATGGCCT GAAATRAACT CTGAAAGAGG AACTTGGTTA GGFACCTATT
 AATAGTATC AATTACGGG TCATTAGTTC ATAGCCATA TATGGAGTTC CGGTTACAT
 AACTTACGGT AAATGGCCCG CCTGGCTGAC CGCCCAACGA CCCCCGCCA TTGACGTCAA
 TAATGACGTA TGTCCCATG GTAACGCCAA TAGGGACTTT CCATTGACGT CAATGGGTGG
 AGTATTTACG GTAAACTGCC CACTTGGCAG TACATCAAGT GTATCATATG CCAAGTACGC
 CCCCATTGA CGTCAATGAC GGTAAATGGC CCGCCTGGCA TTATGCCCG TACATGACCT
 TATGGGACTT TCCTACTGG CAGTACATCT ACGTATTAGT CATCGCTATT ACCATGGTGA
 TCGGGTTTG GCAGTACATC AATGGGCGTG GATAGCGGTT TGACTCACGG GGATTTCCAA
 GTCTCCACC CATTGACGTC AATGGGAGTT TGTTTGGCA CCAAATCAA CGGGACTTTC
 CAAATGTCG TAACAACTCC GCCCAATTGA CGCAAATGGG CGGTAGGCCGT GTACGGTGGG

Figure 13.3

AGGCTATAT AAGCAGAGCT CGTTTAGTGA ACCGTCAGAT CGCCTGGAGA CGCCATCCAC
GCTGTTTTGA CCTGCTAGCT TATCGGGCCG GGAACGGTGC ATTGGAACGC GGATTCCTCCCG
TGCCAAGAGT CAGGTAAGTA CCGCCTATAG AGTCTATAGG CCCACCCCTT TGGCTTCGTT
AGAACGGGC TACAATTAAT ACAATAACCTT TTGGATCGAT CCTACTGACA CTGACATCCA
CTTTTCTTT TTCTCCACAG GTGTCCACTC CCAGTCCAA CTGCACCTCG GTTCGGGAAG
CTCGCTGGG CTGCATCGAT TGAATTCCAC C --Insert Sequence of Interest--
CGATGG CCGCCATGG CCAACTTGT TATTGCAGCT TATAATGGTT
ACAAATAAAG CAATAGCATC ACAAAATTCA CAAATAAAGC ATTTTTTCA CTGCATCTA
GTTTGGTTT GTCCAACTC ATCAATGTAT CTTATCATGT CTGGAICGGG AATTAATTCG
GGCAGCACC ATGGCTGAA ATAAGTTAA ACCCTCTGAA AGAGGAACTT GGTAGGTAC
CGACTAGTCT TTTGCAAAA GCTGTTACCT CGAGCGGCG CTTAATTAAG GCGGCGCATT
TAAATCCTGC AGGTACAGC TTGGCACTGG CCGTCGTTTT ACAACGTCGT GACTGGGAAA
ACCCCTGGGT TACCCAACTT AATCGCCTTG CAGCACATCC CCTTTTCGCC AGCTGGCGTA
ATAGCGAAGA GGCCCGCACC GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT
GGGGCCTGAT GGGGTATTT CTCCTTACGC ATCTGTGGG TATTTCACAC CGCATAACGTC
AAAGCACCA TAGTACGCG CCTGTAGCGG CGCATTAAAC GCGGCGGGTG TGGTGGTTAC
GGCAGCGTG ACCGCTACAC TTGCCAGCG CCTAGCGCC GCTCCTTTCG CTTTCTTCCC
TTCTTCTC GCCACGTTG CCGGCTTTC CCGTCAAGCT CTAATCGGG GGCTCCCTTT

Figure 13.4

AGGGTTCCGA TTTAGTGTCTT TACGGCACCT CGACCCCAA AACTTGATT TGGGTGATGG
TTCACGTAAGT GGGCCAATCGC CCTGATAGAC GGTTTTTCGC CCTTTGACGT TGGAGTCCAC
GTTCTTTAAT AGTGGACICTT TGTTCRAAC TGAACAACA CTCRAACCCTA TCTCGGGCTA
TTCTTTTGAT TTATAAGGA TTTTGCCGAT TTCGGCCTAT TGGTTAAAA ATGAGCTGAT
TTAACAAAA TTTAACCGA ATTTTACAA AATATTAACG TTTACAATTT TATGGTGCAC
TCTCAGTACA ATCTGCTCTG ATGCCGCATA GTTAAGCCAG CCCCAGACACC GCCCCGACAC
CCGCCAACAC CCGCTGACGC GCCCTGACGG GCTTGTCTGC TCCCGGCATC CGCTTACAGA
CAAGTGTGA CCGTCTCCGG GAGCTGCATG TGTACAGAGT TTTACCCCTC ATCACCAGAA
CGCGCGAGAG ACGAAAGGC CTCGTGTATC GCCTATTTTT ATAGTTAAT GTCATGATAA
TAATGGTTTC TTAGACGTA GGTGGCACTT TTCGGGAAA TGTGCGCGGA ACCCCATTTT
GTTTATTTTT CTAATACAT TCAAATATGT ATCCGCTCAT GAGACATAA CCCTGATAAA
TGCTTCAATA ATATTGAAA AGGAGAGTA TGAGTATTCA ACATTTCCGT GTCGCCCTTA
TTCCTTTTT TGCGCATTT TGCCCTCCTG TTTTTGCTA CCCAGAAACG CTGGTAAAG
TAAAGATGC TGAAGATCAG TTGGGTGCAC GAGTGGTTA CATCGAACTG GATCICAACA
GCGGTAAGAT CCTGAGAGT TTTCCGCCCG AAGAAGTTT TCCAATGATG AGCACTTTTA
AAGTCTGCT ATGTGGCGG GTATTATCCC GTATTGACGC CGGCAAGAG CAACTCGGTG
GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA GAAAGCATC
TTACGGATGG CATGACAGTA AGAATATAT GCAGTCTGTC CATACCATG AGTGATAACA

Figure 13.5

CTCGGCCAA CTTACTTCTG ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC
ACACATGGG GGATCAITGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
TACCAACGA CGAGCGTAGC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAC
TATTAATGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC TGGATGGAGG
CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC GGCTGGCTGG TTTATTGCTG
ATAAATCTGG AGCCGGTAGC CGTGGGTCTC GCGGTATGAT TGCAGCACTG GGGCCAGATG
GTAAGCCCTC CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAA
GAAATAGACA GATCGCTGAG ATAGGTGCT CACTGATTAA GCATTGGTAA CTGTCAGACC
AAGTTTACTC ATATAACTT TAGATTGATT TAAAACITCA TTTTAAITT AAAAGGATCT
AGGTGAGAT CCTTTTTGAT AATCTCAAGA CCAAAATCCC TTAACGTGAG TTTTCGTTCC
ACTGAGCTC AGACCCCGTA GAAAAGATCA AAGGATCTC TTGAGATCCT TTTTTCITGC
GCGTAATCTG CTGCTTGCAA ACAAAAAAC CACCGCTACC AGCGGTGGT TGTITGCCGG
ATCAAGAGCT ACCAATCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
ATACTGTTCT TCTAGTGTAG CCGTAGTTAG GCCACCACIT CAAGAACTCT GTAGCACCGC
CTACATACT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC TGCCAGTGGC GATAAGTCGT
GTCTTACCGG GTTGGACTCA AGACGATAGT TACCGGATTA GGGCAGCGG TCGGGCTGAA
CGGGGGTTC GTGCACACAG CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC
TACAGCGTGA GCTATGAGAA AGGCCACCG TTCCCGAAGG GAGAAAGGCG GACAGGTATC

Figure 13.6

CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG GGAACGCCT
GGTAUCTTTA TAGTCTGTG GGGTTTCGCC ACCTCTGACT TGAGCGTCTGA TTTTGTGAT
GCTCGTCAGG GGGCGGAGC CTATGAAAA AGCCAGCAA CGGGCCCTTT TTACGGTTCC
TGGCCTTTTG CTGGCCTTTT GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG
ATAACCGTAT TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
GCAGCGATC AGTGAGCGAG GAAGCGGAG AGGCCCAAT ACGCAAACCG CCTCTCCCCG
CGCGTTGCC GATTCAITAA TGCAGCTGGC ACGACAGGT TCCCGACTGG AAAGCGGGCA
GTGAGCGCAA CGCAATTAAT GTGAGTTAGC TCACTCAITA GGCACCCAG GCTTTACT
TTATGCTTCC GGCTCGTAG TTGTGTGAA TTGTGAGCGG ATAACAATTT CACACAGGAA
ACAGCTATGA CATGATTAGC AATTA

Figure 14. Plasmid SV40.IPD.Heterologous Polypeptide

```

6      <400>
60      TTCGAGCTCG CCGGACATTG ATTATTGACT AGAGTCGATC GACAGCTGTG GAATGTGTGT
120     CAGTTAGGGT GTGGAAAGTC CCCAGGCTCC CCAGCAGGCA GAAGTATGCA AAGCATGCAT
180     CTCAAATTAGT CAGCAACCAG GTGTGGAAG TCCCCAGGCT CCCCAGCAGG CAGAAGTATG
240     CAAAGCATGC ATCTCAATTA GTCAGCAACC ATAGTCCCGC CCCTAACTCC GCCCATCCCG
300     CCCCTAACTC CGCCAGTTC CGCCCAATCT CCGCCCCATG GCTGACTAAT TTTTTTATT
360     TATGCAGAGG CCGAGGCCGC CTCGGCCTCT GAGCTAATCC AGAAGTAGTG AGGAGGCTTT
420     TTTGGAGGCC TAGGCTTTG CAAAAGCTA GCTTATCCGG CCGGGAACGG TGCATTGGAA
480     CGCGGATTCC CCGTGCCAAG AGTGACGTAA GTACCCGCTA TAGAGCGACT AGTCCACCAT
540     GACCGAGTAC AAGCCACGG TGGCCCTCGC CACCCGCGAC GACGTCCCGC GGGCCGTACG
600     CACCCTCGCC GCCGCTTCG CCGACTACCC CGCCACGGC CACACCCTAG ACCCGGACCG
660     CCACATCGAG CGGTCCACCG AGCTGCAAGA ACTCTTCTC ACGCGCTCG GGCTCGACAT
720     CGGCAAGGTG TGGTCCCGG ACGACGGCGC CGCGGTGGCG GTCTGGACCA CGCCGGGAGAG
780     CGTCGAAGCG GGGCCGTGT TCGCCGAGAT CGGCCCGCGC ATGGCCGAGT TGAGCGGTTT
840     CCGCTGGCC GCGCAGCAAC AGATGGAAG CCTCCTGGG CCGCACCGGC CCAAGGAGCC
900     CGCGTGGTTC CTGGCCACCG TCGGCGTCTC GCCCGACCAC CAGGGCAAGG GTCTGGGCAG

```

Figure 14.1

960 CGCCGTCGTG CTCCCCGGGAG TGGAGGGGGC CGAGCGGGCC GGGGTGCCCC CCTTCCTGGA
1020 GACCTCCGGG CCCCAGCAACC TCCCCCTTCTA CGAGCGGGTC GGCTTCACCG TCACCCGCGGA
1080 CGTCGAGTGC CCGAAGGACC GCGGACCTG GTGCATGACC CGCAAGCCCC GTGCCAACAT
1140 GGTTCGACCA TTGAACTGCA TCGTCGCCGT GTCCCAAAAT ATGGGGATTG GCAAGAACGG
1200 AGACCTACCC TGCCCTCCGC TCAGGAACGC GTTCAAGTAC TTCCAAGAA TGACCACAAC
1260 CTCTTCAGTG GAAGGTAAC AGAATCTGGT GATTAATGGT AGGAAAACCT GGTTCCTCCAT
1320 TCCTGAGAAG AATCGACCTT TAAAGGACAG AATTAATATA GTTCTCAGTA GAGAACTCAA
1380 AGAACCAACA CGAGGAGCTC ATTTTCTTGC CAAAAGTTG GATGATGCCCT TAAGACTTAT
1440 TGAACAACCG GAATTGGCAA GTAAAGTAGA CATGGTTTGG APAGTCGGAG GCAGTTCTGT
1500 TTACCAGGAA GCCATGAATC AACCAGGCCA CCTTAGACTC TTTGTGACAA GGATCATGCA
1560 GGAATTTGAA AGTGACACGT TTTTCCCAGA AATTGATTTG GGGAAATATA AACCTCTCCC
1620 AGAATACCCA GCGTCCCTCT CTGAGGTCCA GGAGGAAAAA GGCATCAAGT ATAAGTTTGA
1680 AGTCTACGAG AAGAAAGACT AACGTTAACT GCTCCCCCTCC TAAAGCTATG CATTTTTATA
1740 AGACCATGGG ACTTTTGTCTG GCTTTAGATC CCGTTGGCTT CGTTAGAACC CAGCTACAAT
1800 TAATACATAA CCTTATGTAT CATAACATA CGATTTAGGT GACACTATAG ATAACATCCA
1860 CTTTGCCTTT CTCTCCACAG GTGTCCACTC CCAGGTCCAA CTGCACCTCG GTTCTATCGA
1920 TTGAATTTCA CC - Insert Sequence of Interest -
CGATGGCC GCCATGGCCC AACTTGTTTA TTGCAGCTTA

Figure 14.2

TAATGGTTAC AAATAAAGCA ATAGCATCAC AAATTCACA AAFAAGCAT TTTTTCAC
GCATCTAGT TGTGGTTTGT CCAAACTCAT CAATGTATCT TAICATGTCT GGAICGGGAA
TTAATTCGGC GCAGCACCAT GGCTGAAAT AACCTCTGAA AGAGAACTT GGTAGGTAC
CTTCTGAGGC GGAAGAACC AGCTGTGGAA TGIGTGTGAG TTAGGGTGTG GAAAGTCCCC
AGGCTCCCCA GCAGGCAGAA GTATGCAAG CATGCATCTC AATTAGTCAG CAACCAGGTG
TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC
AGCAACCATA GTCCCGCCCC TAACTCCGCC CATCCGGCC CTAATCCGC CCAGTCCGC
CCATCTCCG CCCCATGGCT GACTAATTTT TTTTATTTAT GCAGAGGCCG AGGCCGCCTC
GGCTCTGAG CTATTCAGA AGTAGTGAGG AGGCTTTTTT GGAGAGCTT TTGCAAAAAG
CTAGCTTATC CGGCCGGAA CGGTGCATTG GAACGGGAT TCCCGTGCC AAGAGTCAGG
TAAGTACCGC CTAFAGAGTC TATAGGCCCA CCCCTTGGC TTCGTTAGAA CGCGGCTACA
ATTAATACAT AACCTTTTGG ATCGATCCTA CTGACACTGA CATCCACTTT TTCTTTTTCT
CCACAGGTGT CCACTCCCAG GTCCAACCTGC ACCTCGGTTC GCGAAGCTAG CTTGGGCTGC
ATCGATTGAA TTCCACC -Insert Sequence of Interest-
CGATGGCCGC CATGGCCCAA CTGTTTATT GCAGTTATA ATGGTTACAA ATAAAGCAAT
AGCATCACAA ATTTACAAA TAAAGCATTT TTTTCACTGC ATTTAGTGTG TGGTTGTCC
AAACTCATCA ATGTATCTTA TCAITCTCTGG ATCGGAAAT AATTCGGCCG AGCACCATGG
CCTGAATAAA GTTTAAACCC TCTGAAAGAG GAAC TTGGTT AGGTACCAGAC TAGTCTTTTG

Figure 14.3

CAAAAAGCTG TTACCTCGAG CGGCCGCTTA ATTAAGGCGC GCCATTTAAA TCCTGCAGGT
AACAGCTTGG CACTGGCCGT CGTTTTACAA CGTCGTGACT GGGAAAACCC TGGCGTTACC
CAACTTAATC GCCTTGCAGC ACATCCCCCT TTCGCCAGCT GCGGTAATAG CGAAGAGGCC
CGCACCGATC GCCCTTCCCA ACAGTTGCGC AGCCTGAATG GCGAATGGCG CCTGATGCGG
TATTTTCTCC TTACGCCATCT GTGCGGTATT TCACACCGCA TAGGTCAAAG CAACCATAGT
ACGGGCCCTG TAGGGGCGCA TTAAGCGCGG CGGGTGTGGT GGTACGCGC AGCGTGACCG
CTACACTGC CAGGCCCTA GCGCCCGCTC CTTTGGCTTT CTCCCTTCC TTCTCGCCA
CGTTCCGGG CTTTCCCGGT CAAGCTCTAA ATCGGGGGT CCCTTTAGGG TTCCGATTTA
GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTTGGG TGATGGTTCA CGTAGTGGG
CATCGCCCTG ATAGACGGT TTTCGCCCTT TGACGTTGGA GTCCACGTTT TTTAATAGTG
GACTCTTGT CCAAACTGGA ACAACACTCA ACCCTATCTC GGGCTATTCT TTTGATTTAT
AAGGGATTT GCCGATTTG GCCTATTGGT TAAAAAATGA GCTGATTTAA CAAAAATTTA
ACGGGAATTT TAACAAAATA TTAACGTTTA CAATTTTATG GTGCACCTCTC AGTACAATCT
GCTCTGATCC CGCATAGTTA AGCCAGCCCC GACACCCGCC AACACCCGCT GACGGGCCCT
GACGGGCTG TCTGCTCCCG GCATCCGCTT ACAGACAAGC TGTGACCGTC TCCGGGGAGCT
GCATGTGCA GAGGTTTCA CCGTCAACAC CGAAACGGCC GACGAAAGGG CCTCGTGATA
CGCCTATTTT TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT
TTTCCGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA TTCAAATATG

TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT
ATGAGTATC AACATTTCCG TGTGCCCCIT ANTCCCTTTT TTGCGGCATT TTGCCTTCCT
GTTTTGCTC ACCCAGAAC GCTGGTGAAA GTAARAAGATG CTGAAGATCA GTTGGGTGCA
CGAGTGGGT ACATCGACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
GAAGAAGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC GGTATTATCC
CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG
GTTGAGTACT CACCAGTCAC AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA
TGCAGTGTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC
GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA CACCACGATG
CCTGTAGCAA TGGCAACAC GTTGCGCAA CTATTAAC TGCGAACTACT TACTCTAGCT
TCCCGGCAAC AATTAATAGA CTGGATGGAG CCGGATAAAG TTGCAGGACC ACTTCTGCGC
TCGGCCCTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGCT
CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC
ACGACGGGA GTCAGGCAAC TATGGATGAA CGAATAGAC AGATCGCTGA GATAGGTGCC
TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT
TTAAAACCTC ATTTTTAAT TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG
ACCAAAATCC CTTAACGTGA GTTTTCGTTT CACTGAGCGT CAGACCCCGT AGAAAAGATC

AAAGGATCTT CTTGAGATCC TTTTTTCTG CCGGTAATCT GCTGTGTGCA AACAAAAAAA
CCACCCGTAC CAGCGGTGGT TTGTTTGCCG GATCAAGAGC TACCAACTCT TTTTCCGAAG
GTAAC TGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA
GGCCACCACT TCAAGAACTC TGTAGCACCG CCFACATAACC TCGCTCTGCT AATCCTGTTA
CCAGTGGCTG CTGCCAGTGG CGATAAGTGG TGCTTTACCG GGTGGACTC AAGACGATAG
TTACCGGATA AGCCGCAGCG GTCGGGCTGA ACGGGGGGTT CCFACACACA GCCCAGCTTG
GAGCGAACGA CCTACACCGA ACTGAGATAC CTACAGCGTG ACCTATGAGA AAGCGCCACG
CTTCCCGAAG GGAGAAAGCC GGACAGGTAT CCGGTAAGCG GCAGGGTCCG AACAGGAGAG
CGCAGGAGG AGCTTCCAGG GGGAAACGCC TGGTATCTTT ATAGTCCTGT CCGGTTTCGC
CACCTCTGAC TTGAGCGTGC ATTTTGTGA TGCTCGTCAG GGGGGCGGAG CCTATGGAAA
AACGCCAGCA ACGCGGCCTT TTTACGGTTC CTGGCCCTTT GCTGGCCCTT TGCTCACATG
TTCTTTCTG CGTTATCCCC TGATTCTGTG GATAACCGTA TTACCGCCTT TGAGTGAGCT
GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGACT CAGTGAGCGA GGAAGCCGAA
GAGCGCCAA TAGGCAACC GCCTCTCCCC GCGCGTTGGC CGATTCAFTA ATGCAGCTGG
CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTA TGTGAGTTAG
CTCACTCATT AGGCACCCCA GGCCTTACAC TTTAIGCTTC CCGCTCGTAT GTTGTGTGGA
ATTGTGAGCG GATAACAATT TCACACAGGA AACAGTATG ACATGATTAC GAATTA

Figure 14.6

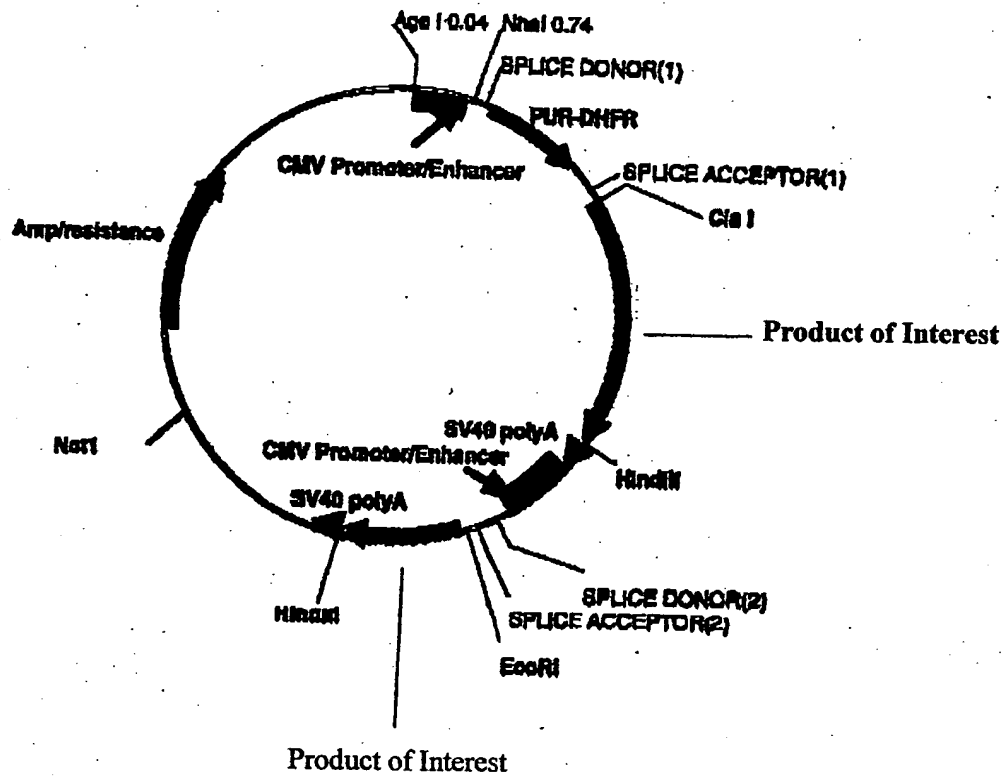


Figure 15. pCMV.IPD.HP

Timeline and Titer Comparison

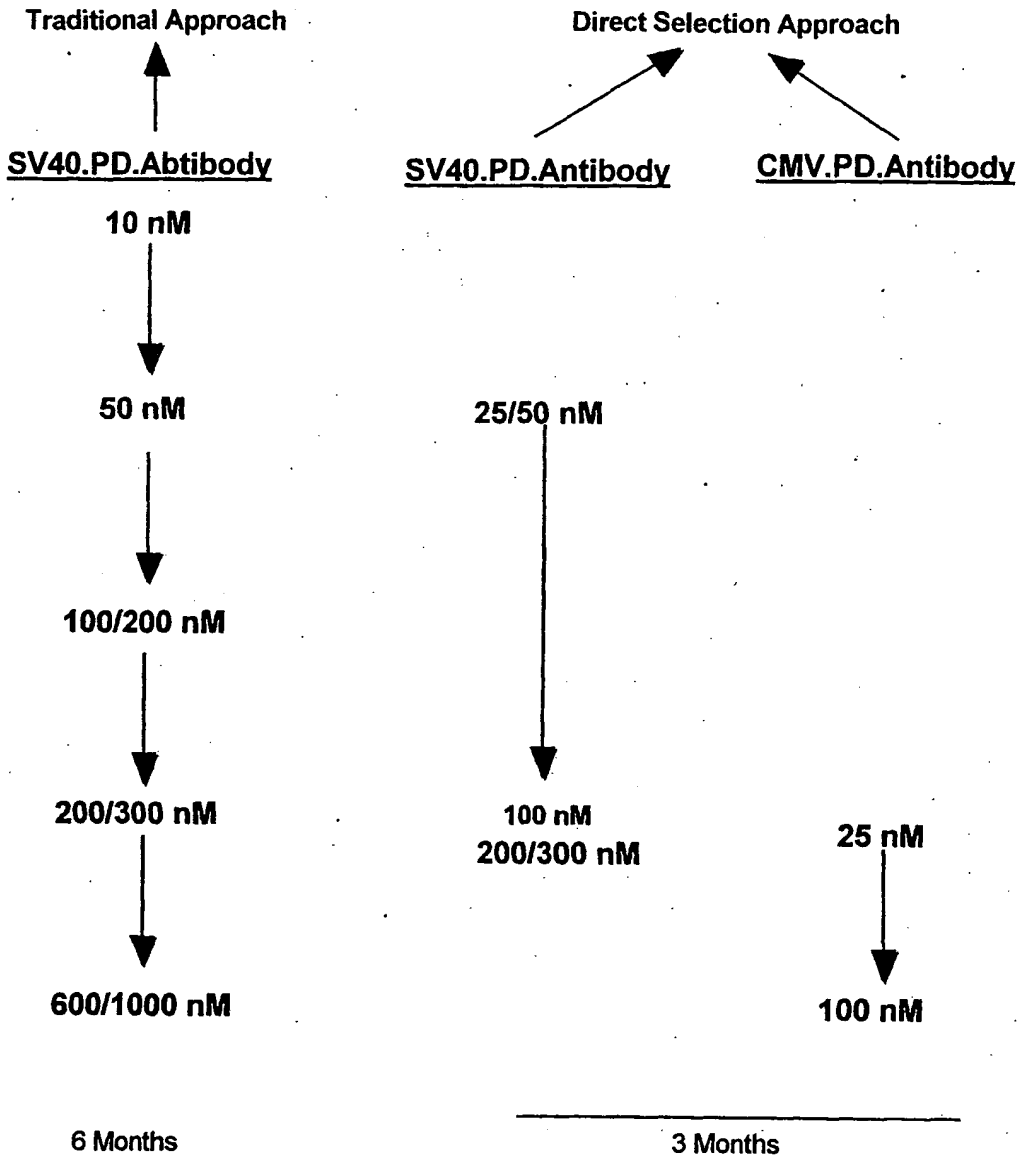


Figure 16. Timeline and Titer Comparison.

SEQUENCE LISTING

<110> Kruppen, Lynne
 Shen, Amy
 Chisum, Venessa

<120> INTRON FUSION CONSTRUCT AND METHOD OF USING FOR SELECTING HIGH-
 EXPRESSING
 PRODUCTION CELL LINES

<130> 22338/00101

<150> US 60/426,095

<151> 2002-11-14

<160> 4

<170> PatentIn version 3.1

<210> 1

<211> 5147

<212> DNA

<213> Artificial

<220>

<223> plasmid pSV.IPUR circular ds-DNA

<400> 1

```

ttcgagctcg cccgacattg attattgact agagtcgac gacagctgtg gaatgtgtgt      60
cagttagggg gtggaaagtc ccagggctcc ccagcaggca gaagtatgca aagcatgcat      120
ctcaattagt cagcaaccag gtgtggaaag tccccaggct cccagcagg cagaagtatg      180
caaagcatgc atctcaatta gtcagcaacc atagtcccgc ccctaactcc gcccatcccg      240
cccctaactc cgcccagttc cgcccattct cggcccatg gctgactaat tttttttatt      300
tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt      360
tttgagggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa      420
cgcggattcc ccgtgccaaag agtgacgtaa gtaccgccta tagagcgact agtccaccat      480
gaccgagtac aagcccacgg tgcgcctcgc caccgcgcac gacgtcccc gggccgtacg      540
caccctcgcc gcccggttcg ccgactacct cgccacgcgc cacaccgtcg acccggaccg      600
ccacatcgag cgggtcaccg agctgcaaga actcttctc acgcgcgtcg ggctcgacat      660
cggcaagggt tgggtcgcgg acgacggcgc cgcggtggcg gtctggacca cgccggagag      720
cgtcgaagcg ggggcggtgt tcgccgagat cggcccgcgc atggccgagt tgagcggttc      780
ccggtggcc gcgcagcaac agatggaagg cctctggcg ccgcaccggc ccaaggagcc      840
cgcgtggttc ctggccaccg tcggcgtctc gcccgaccac cagggaagg gtctgggcag      900

```

cgccgtcgtg ctccccggag tggaggcggc cgagcgcgcc ggggtgcccg ccttcctgga 960
gacctccgcg ccccgcaacc tccccttcta cgagcggctc ggcttcaccg tcaccgccga 1020
cgtcagtgtc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagcccg gtgcctgagt 1080
taactgctcc cctcctaaag ctatgcattt ttataagacc atgggacttt tgetggcttt 1140
agatcccctt ggcttcgta gaacgcagct acaattaata cataacctta tgtatcatac 1200
acatacgatt taggtgacac tatagataac atccactttg cttttctctc cacaggtgtc 1260
cactcccagg tccaactgca cctcggttct atcgattgaa ttccccgggg atcctctaga 1320
gtcgaacctg agaagcttgc atggccgcca tggeccaact tgtttattgc agcttataat 1380
ggttacaat aaagcaatag catcacaat ttcacaata aagcattttt ttcactgcat 1440
tctagttgtg gtttgccaa actcatcaat gtatcttata atgtctggat cgatcgggaa 1500
ttaattcggc gcagcaccat ggcctgaaat aacctctgaa agaggaactt ggtaggtac 1560
cttctgaggc ggaaagaacc agctgtggaa tgtgtgtcag ttaggggtgtg gaaagtcccc 1620
aggctcccca gcagcgagaa gtatgcaaag catgcatctc aattagtcag caaccaggtg 1680
tggaaagtcc ccaggctccc cagcaggcag aagtatgcaa agcatgcatc tcaattagtc 1740
agcaaccata gtcccccccc taactccgcc catccccccc ctaactccgc ccagttccgc 1800
ccattctccg ccccatggct gactaatttt tttatttat gcagaggccg aggccgcctc 1860
ggcctctgag ctattccaga agtagtgagg aggccttttt ggaggcctag gcttttgcaa 1920
aaagctgtta cctcgagcgg ccgcttaatt aaggcgcgcc atttaaacc tgcaggtaac 1980
agcttggcac tggccgctgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 2040
cttaatcgcc ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccgc 2100
accgatcgcc ctccccaaaca gttgcgtagc ctgaatggcg aatggcgctt gatgcgggat 2160
tttctcctta cgcactctgt cggatattca caccgcatac gtcaaagcaa ccatagtacg 2220
cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc gtgaccgcta 2280
cacttgccag cgccttagcg cccgctcctt tcgctttctt ccttccttt ctgcgccagt 2340
tcgccggctt tccccgtaa gctctaaatc gggggctccc tttagggttc cgatttagtg 2400
ctttaoggca cctcgacccc aaaaaacttg atttgggtga tggttcacgt agtgggccat 2460
cgccctgata gacggttttt cgcccttga cgttggagtc cacgttcttt aatagtgagc 2520
tcttgttcca aactggaaca aactcaacc ctatctcggg ctattctttt gatttataag 2580

ggatTTTgCC gattTCggcc tattgGTTaa aaaatgagct gattTaaCaA aaattTaaCg 2640
cgaattTTTaa caaaatatta acgTTTaaCa tTTtatggtg cactctcagt acaatctgct 2700
ctgatgccgc atagTTaaGc caactccgct atcgctacgt gactgggtca tggctgcgcc 2760
ccgacaccCG ccaacaccCG ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc 2820
ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt cagaggTTTT caccgtcatc 2880
accgaaacgc gcgaggcagt attcttgaag acgaaagggc ctcgtgatac gcctatTTTT 2940
ataggttaat gTcatgataa taatggttTc ttagacgtca ggtggcactt ttcggggaaa 3000
tgtgcgcgga acccctatTT gTTtatTTTT cTaaatacat tCaaatatgt atccgctcat 3060
gagacaataa ccctgataaa tgcttcaata atattgaaa aggaagagta tgagtattca 3120
acatttccgt gTcgcctta tTccTTTT tgcggcattt tgccttcotg tTTTtGctca 3180
cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggTta 3240
catcgaactg gatctcaaca gcggtaaGat ccttgagagt tTtcgccccg aagaacgTtt 3300
tccaatgatg agcactTTTta aagttctgct atgtggcgcg gtattatccc gtgatgacgc 3360
cgggcaagag caactcggTc gcgcataca ctattctcag aatgactTgg ttgagtactc 3420
accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgtctg 3480
cataaccatg agtgataaca ctgcggccaa ctTacttctg acaacgatcg gaggaccgaa 3540
ggagctaacc gctTTTTtgc acaacatggg ggatcatgta actcgcctTg atcgTtggga 3600
accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc cagcagcaat 3660
ggcaacaacg ttgcgcaaac tattaactgg cgaactactt actctagctt cccggcaaca 3720
attaatagac tggatggagg cggataaagt tgcaggacca ctTctgcgct cggccctTcc 3780
ggctggctgg tTtattgctg ataaatctgg agccggtgag cgtgggtctc gcggtatcat 3840
tgacgactg gggccagatg gtaagccctc ccgtatcgta gttatctaca cgacggggag 3900
tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa 3960
gcattggtaa ctgtcagacc aagTTTactc atataactt tagattgatt taaaactTca 4020
TTTTtaattt aaaaggatct aggtgaagat cTTTTtgat aatctcatga cCaaaatccc 4080
ttaacgtgag tTTtGttcc actgagcgtc agaccCCgta gaaaagatca aaggatcttc 4140
ttgagatcct tTTTTtctgc gcgTaatctg ctgcttgcaa aCaaaaaac caccgctacc 4200
agcggTggtt tGTTTTgcgg atcaagagct accaactctt tTccgaagg taactggctt 4260
cagcagagcg cagatacCaA atactgtcct tctagtgtag ccgtagttag gccaccactt 4320

```

caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc 4380
tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa 4440
ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg agcgaacgac 4500
ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcgccaacgc ttcccgaagg 4560
gagaaaggcg gacaggtatc cggtaagcgg cagggtcgga acaggagagc gcacgagggg 4620
gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 4680
tgagcgtcga tttttgtgat gctcgtcagg gggcgggagc ctatggaaaa acgccagcaa 4740
cgcggccttt ttaacggttc tggccttttg ctggcctttt gtcacatgt tcttctctgc 4800
gttatcccct gattctgtgg ataaccgtat taccgccttt gaggagctg ataccgctcg 4860
ccgcagccga acgaccgagc gcagcagtc agtgagcagag gaagcgggaag agcgcccaat 4920
acgcaaaccg cctctccccg cgcgttggcc gattcattaa tccagctggc acgacaggtt 4980
tcccgactgg aaagcgggca gtgagcgcga cgcaattaat gtgagttacc tcactcatta 5040
ggcaccocag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg 5100
tattgttaaa gtgtgtcctt tgtcgatact ggtactaatg ctttaatt 5147

```

```

<210> 2
<211> 5171
<212> DNA
<213> Artificial

```

```

<220>
<223> plasmid pSV.ID circular ds-DNA

```

```

<220>
<221> misc_feature
<222> (444)..(444)
<223> splice donor

```

```

<220>
<221> misc_feature
<222> (1946)..(1946)
<223> start pUC118

```

```

<220>
<221> misc_feature
<222> (1954)..(1954)
<223> linearization linker inserted into HpaI site

```

```

<220>

```

<221> misc_feature
 <222> (529)..(1090)
 <223> DHFR coding region

<220>
 <221> misc_feature
 <222> (1522)..(1522)
 <223> sv40 origin

<400> 2
 ttcgagctcg cccgacattg attattgact agagtcgatc gacagctgtg gaatgtgtgt 60
 cagttagggg gtggaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
 ctcaattagt cagcaaccag ggtggaaag tccccaggct cccagcagg cagaagtatg 180
 caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gccatcccg 240
 ccctaactc cgcccagttc cgcccattct cggcccotg gctgactaat tttttttatt 300
 tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt 360
 tttggaggcc taggcttttg caaaaagcta gcttatccgg ccgggaacgg tgcattggaa 420
 cgcgattcc ccgtgccaa agtgacgtaa gtaccgcta tagagtctat aggcccacc 480
 cttggctcta gagagatata agcctaggat tttatcccc gtgccatcat ggttcgacca 540
 ttgaactgca tcgtcgccgt gtcccaaat atggggattg gcaagaacgg agacctacc 600
 tgccctccgc tcaggaacgc gttcaagtac ttcaaagaa tgaccacaac ctcttcagt 660
 gaaggtaaac agaatctggt gattatgggt aggaaaacct ggttctccat tcttgagaag 720
 aatcgacct taaaggacag aattaatata gttctcagta gagaactcaa agaaccacca 780
 cgaggagctc attttcttg caaaagtttg gatgatgcct taagacttat tgaacaaccg 840
 gaattggcaa gtaaagtaga catggtttgg atagtcggag gcagttctgt ttaccaggaa 900
 gccatgaatc aaccaggcca ccttagactc tttgtgaaa ggatcatgca ggaatttgaa 960
 agtgacacgt ttttccaga aattgatttg gggaaatata aacctctccc agaataccca 1020
 ggcgtcctct ctgaggtcca ggaggaaaa ggcacaaagt ataagtttga agtctacgag 1080
 aagaaagact aacaggaaga tgctttcaag ttctctgctc ccctoctaaa gctatgcatt 1140
 tttataagac catgggactt ttgctggctt tagaccctct tggcttcggt agaacgcggc 1200
 tacaattaat acataacctt atgtatcata cacatagatt taggtgacac tatagaataa 1260
 catccacttt gcttttctct ccacaggtgt cactccaggt caactgcacc tcggttctat 1320
 cgattgaatt ccccggggat cctctagagt cgacctgcag aagcttggcc gccatggccc 1380

aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac aaatttcaca 1440
aataaagcat ttttttcaact gcattctagt tgtggtttgt ccaaactcat caatgtatct 1500
tatcatgtct ggatcgatcg ggaattaatt cggcgcagca ccatggcctg aaataacctc 1560
tgaaagagga acttggttag gtaccttctg aggcggaaag aaccagctgt ggaatgtgtg 1620
tcagttaggg tgtggaaagt cccaggctc cccagcaggc agaagtatgc aaagcatgca 1680
tctcaattag tcagcaacca ggtgtggaaa gtccccaggc tccccagcag gcagaagtat 1740
gcaaagcatg catctcaatt agtcagcaac catagtcccg cccctaactc cgcccatccc 1800
gccctaact ccgcccagtt ccgcccattc tccgcccatt ggctgactaa ttttttttat 1860
ttatgcagag gccgaggccg cctcggcctc tgagctattc cagaagtagt gaggaggctt 1920
ttttggaggc ctaggctttt gcaaaaagct gttacctoga gggccgctt aattaaggcg 1980
cgccatttaa atcctgcagg taacagcttg gactggccg tcgttttaca acgtcgtgac 2040
tgggaaaacc ctggcgttac ccaacttaat cgccttgag cacatcccc cttcgccagc 2100
tggcgtaata gcgaagaggc ccgcaccgat cgccttccc aacagttgag tagcctgaat 2160
ggcgaatggc gcctgatgag gtattttctc cttacgcac tgtgcggtat ttcacaccgc 2220
atacgtcaaa gcaaccatag tacgcgcctt gttagcggcgc attaagcgcg gcgggtgtgg 2280
tggttacgag cagcgtgacc gctacacttg ccagcgcctt agcgcgccgt cttttcgtt 2340
tcttccctc ctttctgcc acgttcgccg gctttcccc tcaagctcta aatcgggggc 2400
tccttttagg gttccgattt agtgctttac ggcacctoga ccccaaaaaa cttgatttgg 2460
gtgatggttc acgtagtggg ccacgcctt gatagacggt ttttcgcctt ttgacgttgg 2520
agtccacggt ctttaatagt ggactcttgt tccaaactgg aacaacactc aaccctatct 2580
cgggctatc ttttgattta taagggattt tgccgatttc ggcctattgg ttaaaaaatg 2640
agctgattta acaaaaattt aacgcgaatt ttaacaaaat attaacgttt acaattttat 2700
ggtgactct cagtacaatc tgctctgatg ccgcatagtt aagccaactc cgctatcgtc 2760
acgtgactgg gtcattggctg cgcgccgaca cccgccaaca cccgctgacg cgcctgacg 2820
ggcttgtctg ctcccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat 2880
gtgtcagagg ttttcaccgt catcaccgaa acgcgcgagg cagtattctt gaagacgaaa 2940
gggcctcgtg atacgcctat ttttataggt taatgtcatg ataataatgg tttcttagac 3000
gtcagggtggc acttttcggg gaaatgtgag cggaaaccct atttgtttat ttttctaaat 3060

acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc aataatattg 3120
aaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct tttttgcggc 3180
atthtgcctt cctgthtttg ctcaaccaga aacgtgtgtg aaagtaaaag atgctgaaga 3240
tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta agatccttga 3300
gagthttcgc cccgaagaac gthttccaat gatgagcaat tttaaagttc tgctatgtgg 3360
cgcggtatta tcccgatgag acgcccggca agagcaactc ggtcgccgca tacactattc 3420
tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg atggcatgac 3480
agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg ccaacttact 3540
tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca tgggggatca 3600
tgtaactcgc cttgatcgtt ggaaccgga gctgaatgaa gccataccea acgacgagcg 3660
tgacaccacg atgccagcag caatggcaac aacgttgccg aaactattaa ctggcgaact 3720
acttactcta gcttcccggc aacaattaat agactggatg gaggcggata aagttgcagg 3780
accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat ctggagccgg 3840
tgagcgtggg tctcgcggta tcattgcagc actggggcca gatggttaagc cctcccgtat 3900
cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata gacagatcgc 3960
tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt actcatatat 4020
actttagatt gatttaaaac ttcatthtta atttaaagg atctaggtga agatcctttt 4080
tgataatctc atgacaaaaa tcccttaacg tgagthttcg ttccactgag cgtcagacct 4140
cgtagaaaag atcaaaggat cttcttgaga tctthttttt ctgcgcgtaa tctgctgctt 4200
gcaaacaaaa aaaccaccgc taccagcggg gthttgthtt ccggatcaag agctaccaac 4260
tctthttccg aaggttaactg gcttcagcag agcgcagata ccaaatactg tcttctagt 4320
gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat acctegctct 4380
gctaatacctg ttaccagtgg ctgctgccag tggcgataag tctgtctta ccgggttga 4440
ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg gttcgtgcac 4500
acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc gtgagcattg 4560
agaaagcgc acgcttcccg aaggagaaa ggcggacagg tatccggtaa gcggcaggt 4620
cggaacagga gagcgcacga gggagcttcc aggggaaac gcctgtatc tttatagtcc 4680
tgtcgggtt cgccacctct gacttgagcg tcatthttg tgatgctctg cagggggcg 4740
gagcctatgg aaaaacgcca gcaacgcggc cthtttacgg ttctggcct tttgctggcc 4800

ttttgctcac atgttctttc ctgcgttata ccttgattct gtggataacc gtattaccgc 4860
 ctttgagtga gctgataccg ctgcgccgag ccgaacgacc gagcgcagcg agtcagtgag 4920
 cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcggt ggccgattca 4980
 ttaatccagc tggcagcaca ggtttcccga ctggaaagcg ggcagtgagc gcaacgcaat 5040
 taatgtgagt tacctcactc attaggcacc ccaggcttta cactttatgc ttccggctcg 5100
 tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct atgaccatga 5160
 ttacgaatta a 5171

<210> 3
 <211> 5712
 <212> DNA
 <213> Artificial

<220>
 <223> plasmid pSV.IPD circular ds-DNA

<220>
 <221> misc_feature
 <222> (444)..(444)
 <223> splice donor

<220>
 <221> misc_feature
 <222> (479)..(479)
 <223> start PUR coding

<220>
 <221> misc_feature
 <222> (1079)..(1643)
 <223> DHFR coding region

<400> 3
 ttcgagctcg cccgacattg attattgact agagtcgatc gacagctgtg gaatgtgtgt 60
 cagttagggt gtggaaagtc ccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
 ctcaattagt cagcaaccag gtgtggaaag tccccaggct ccccagcagg cagaagtatg 180
 caaagcatgc atctcaatta gtcagcaacc atagtccgc ccctaactcc gcccatcccg 240
 ccctaactc cgcccagttc cgcccattct ccgccccatg gctgactaat tttttttatt 300
 tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt 360
 tttggaggcc taggcttttg caaaaagcta gttatccgg ccgggaacgg tgcattggaa 420

cgcgattcc ccgtgccaag agtgacgtaa gtaccgccta tagagcgact agtccacat 480
 gaccgagtac aagcccacgg tgcgcctcgc caccgcgac gacgtcccgc gggcgtacg 540
 caccctcgcc gccggttcg ccgactacc ccgcacgcgc cacaccgtag acccgaccg 600
 ccacatcgag cgggtcaccg agctgcaaga actcttctc acgcgcgtcg ggctcgacat 660
 cggcaagtg tgggtcgcg acgacggcgc cgcggtggcg gtctggacca gcgggagag 720
 cgctgaagcg ggggcggtgt tcgcccagat cggcccgcgc atggccgagt tgagcggttc 780
 ccggctggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc 840
 cgcgtggttc ctggccaccg tcggcgtctc gcccgaccac cagggcaagg gtctggcag 900
 cgccgtcgtg ctcccggag tggaggcggc cgagcgcgc ggggtgccc ccttctgga 960
 gacctcgcg ccccgcaacc tccccttcta cgagcggctc ggcttcaccg tcaccgccga 1020
 cgtcgagtgc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagccc gtgccaacat 1080
 ggttcgacca ttgaactgca tcgtcgcgt gtccaaaat atggggattg gcaagaacgg 1140
 agacctacc tgcctcgc tcaggaacgc gttcaagtac ttccaaagaa tgaccacaac 1200
 ctcttcagt gaagtaaac agaactcgt gattatgggt aggaaaacct ggttctccat 1260
 tcctgagaag aatcgacct taaaggacag aattaatata gttctcagta gagaactcaa 1320
 agaaccacca cgaggagctc attttcttgc caaaagttg gatgatgcct taagacttat 1380
 tgaacaaccg gaattggcaa gtaaagtaga catggtttg atagtcggag gcagttctgt 1440
 ttaccaggaa gccatgaatc aaccaggcca ccttagactc tttgtgaaa ggatcatgca 1500
 ggaattgaa agtgacacgt ttttccaga aattgattg gggaaatata aacctctccc 1560
 agaataccca ggcgtcctct ctgaggcca ggaggaaaaa ggcatcaagt ataagttga 1620
 agtctacgag aagaaagact aacgttaact gctcccctcc taaagctatg cattttata 1680
 agaccatgg acttttgctg gcttagatc ccctggctt cgttagaacg cagctacaat 1740
 taatacataa ccttatgtat catacacata cgatttaggt gacactatag ataacatcca 1800
 ctttgccttt ctctccacag gtgtccactc ccaggcca ctgcacctc gttctatcga 1860
 ttgaattccc cgggatcct ctagagtoga cctgcagaag cttcgatggc cgccatggcc 1920
 caacttggtt attgcagctt ataatggtta caataaagc aatagcatca caaattcac 1980
 aaataaagca ttttttcac tgcattctag ttgtggttg tccaaactca tcaatgtatc 2040
 ttatcatgtc tggatcgtc gggaattaat tcggcgcagc accatggcct gaataacct 2100
 ctgaaaggg aacttggtta ggtaccttct gaggcgaaa gaaccagctg tggaatgtgt 2160

gtcagttagg gtgtggaag tccccaggct ccccagcagg cagaagtatg caaagcatgc 2220
 atctcaatta gtcagcaacc aggtgtggaa agtccccagg ctccccagca ggcagaagta 2280
 tgcaaagcat gcatctcaat tagtcagcaa ccatagtccc gccccctaact ccgcccatac 2340
 cgccccctaac tccgccaggt tccgccatt ctccgcccc tggctgacta atttttttta 2400
 tttatgcaga ggccgaggcc gcctcggcct ctgagctatt ccagaagtag tgaggaggct 2460
 tttttggagg cctaggcttt tgcaaaaagc tgttacctcg agcggccgct taattaaggc 2520
 ggcatttta aatcctgcag gtaacagctt ggactggcc gtcgttttac aacgtcgtga 2580
 ctgggaaaac cctggcggtta cccaacttaa tcgccttgca gcacatcccc ccttcgccag 2640
 ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc gttagcctgaa 2700
 tggcgaatgg cgcctgatgc ggtatcttct ccttacgcat ctgtgcggta tttcacaccg 2760
 catacgtcaa agcaaccata gtacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg 2820
 gtggttacgc gcagcgtgac cgctacactt gccagcgcoc tagcgcocgc tcctttcgtc 2880
 ttcttcctt cctttctcgc cacgttcgcc ggctttcccc gtcaagctct aaatcggggg 2940
 ctccctttag ggttcgatt tagtgcttta cggcacctcg acccaaaaa acttgattg 3000
 ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg ttttcgccc tttgacgttg 3060
 gagtccacgt tcttaaatag tggactcttg ttccaaactg gaacaacact caaccctatc 3120
 tcgggctatt cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat 3180
 gagctgattt acaaaaatt taacgcgaat ttaacaaaa tattaacgtt tacaatttta 3240
 tggtgactc tcagtacaat ctgctctgat gccgcatagt taagccaact ccgctatcgc 3300
 tacgtgactg ggtcatggct gcgccccgac acccgccaac acccgctgac gcgccctgac 3360
 gggcttgtct gctcccggca tccgcttaca gacaagctgt gaccgtctcc gggagctgca 3420
 tgtgtcagag gttttaccg tcatcacoga aacgcgcgag gcagtattct tgaagacgaa 3480
 agggcctcgt gatagccta tttttatagg ttaatgtcat gataataatg gtttcttaga 3540
 cgtcaggtag cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 3600
 tacattcaaa tatgtatccg ctcatgagac aataaccctg ataatgctt caataatatt 3660
 gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc tttttgagg 3720
 cttttgctc tcctgttttt gctcaccag aaacgctggg gaaagtaaaa gatgctgaag 3780
 atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg 3840

agagttttcg ccccaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg 3900
gcgcgggtatt atcccgtgat gacgccgggc aagagcaact cggtcgccgc atacactatt 3960
ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga 4020
cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac 4080
ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc 4140
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 4200
gtgacaccac gatgccagca gcaatggcaa caacgttgcg caaactatta actggcgaac 4260
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgacg 4320
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg 4380
gtgagcgtgg gtctcgcggg atcattgcag cactggggcc agatggtaag ccctcccgta 4440
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg 4500
ctgagatagtg tgccctactg attaagcatt ggtaactgtc agaccaagtt tactcatata 4560
tacttttagat tgatttaaaa cttcattttt aatttaaag gatctaggtg aagatccttt 4620
ttgataatct catgacaaaa atcccctaac gtgagttttc gttccactga gcgtcagacc 4680
ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct 4740
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa 4800
ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag 4860
tgtagccgta gttagccac cacttcaaga actctgtagc accgcctaca tacctcgtc 4920
tgctaactct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttg 4980
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 5040
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagcatt 5100
gagaaagcgc cacgcttccc gaaggagaa aggcggacag gtatccggtg agcggcaggg 5160
tcggaacagg agagcgacg agggagcttc cagggggaaa cgcttggtat ctttatagtc 5220
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 5280
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc 5340
cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg 5400
cctttgagtg agctgatacc gctcgcgca gccgaacgac cgagcgcagc gagtcagtga 5460
gcgaggaagc ggaagagcgc ccaatacga aaccgcctct ccccgcgctg tggccgattc 5520
attaatccag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 5580

ttaatgtgag ttacctcaact cattaggcac cccaggcttt acactttatg cttccggctc 5640
 gtatgttggtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 5700
 attacgaatt aa 5712

<210> 4
 <211> 12514
 <212> DNA
 <213> Artificial

 <220>
 <223> plasmid pSV.IPD.2C4 circular ds-DNA

 <220>
 <221> misc_feature
 <222> (444)..(444)
 <223> splice donor

<220>
 <221> misc_feature
 <222> (479)..(479)
 <223> start PUR coding

<220>
 <221> misc_feature
 <222> (1079)..(1643)
 <223> DHFR coding region

<220>
 <221> misc_feature
 <222> (1883)..(1883)
 <223> start 2C4 HC coding

<220>
 <221> misc_feature
 <222> (4154)..(4154)
 <223> start LC coding

<400> 4
 ttcgagctcg cccgacattg attattgact agagtcgacgac gacagctgtg gaatgtgtgt 60
 cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca aagcatgcat 120
 ctcaattagt cagcaaccag gtgtggaaag tccccaggct ccccagcagg cagaagtatg 180
 caaagcatgc atctcaatta gtcagcaacc atagtcccgc ccctaactcc gcccatcccg 240
 ccctaactc cgcccagttc cgcccattct cgcgcccatg gctgactaat tttttttatt 300

tatgcagagg ccgaggccgc ctggcctct gagctattcc agaagtagtg aggaggcttt 360
tttgagggcc taggcttttg caaaaagcta gcttatccgg ccggaacgg tgcatggaa 420
cgcgattcc ccgtgccaag agtgacgtaa gtaccgcta tagagcgact agtccacat 480
gaccgagtac aagcccacgg tggcctcgc caccgcgac gacgtccgc gggccgtacg 540
caccctcgc gccgcgttc cgactacce cgccacgcgc cacaccgtag acccggaccg 600
ccacatcgag cgggtcaccg agctgcaaga actcttctc acgcgcgtcg ggctcgacat 660
cggcaagggtg tgggtcgcgg acgacggcgc cgcggtggcg gtctggacca cgccggagag 720
cgtcgaagcg ggggcggtgt tcgccgagat cgcccggc atggccgagt tgagcggttc 780
ccggctggcc gcgcagcaac agatggaagg cctcctggcg ccgcaccggc ccaaggagcc 840
cgcgtggttc ctggccaccg tcggcgtctc gcccgaccac cagggcaagg gtctgggcag 900
cgcogtcgtg ctccccggag tggaggcggc cgagcgcgc ggggtgcccg ccttctgga 960
gacctccgc ccccgcaacc tccccttcta cgagcgctc ggcttcaccg tcaccgccga 1020
cgtcgagtgc ccgaaggacc gcgcgacctg gtgcatgacc cgcaagcccg gtgccaacat 1080
ggttcgacca ttgaactgca tcgtcgccgt gtccaaaat atggggattg gcaagaacgg 1140
agacctacc tgccctccgc tcaggaacgc gttcaagtac ttcaaagaa tgaccacaac 1200
ctcttcagtg gaaggtaaac agaactctgt gattatgggt aggaaaacct ggttctccat 1260
tcctgagaag aatcgacctt taaaggacag aattaatata gttctcagta gagaactcaa 1320
agaaccacca cgaggagctc atttcttgc caaaagttg gatgatgcct taagacttat 1380
tgaacaaccg gaattggcaa gtaaagtaga catggtttg atagtcggag gcagttctgt 1440
ttaccaggaa gccatgaatc aaccaggcca ccttagactc tttgtgaaa ggatcatgca 1500
ggaatttgaa agtgacacgt tttcccaga aattgattg gggaaatata aacctctccc 1560
agaataocca ggcgtcctct ctgaggtcca ggaggaaaa ggcatcaagt ataagtttga 1620
agtctacgag aagaaagact aacgttaact gtcacctcc taaagctatg ctttttata 1680
agaccatggg acttttgctg gctttagatc cccttgctt ogttagaacg cagctacaat 1740
taatacataa ccttatgtat catacacata cgatttaggt gacactatag aataacatcc 1800
actttgctt tctctccaca ggtgtccact ccaggtcca actgcacctc gttctatcg 1860
attgaattcc accatgggat ggtcatgtat catcctttt ctagtagcaa ctgcaactgg 1920
agtacattca gaagttcagc tgggtgagtc tggcgggtgc ctggtgcagc cagggggctc 1980

actccgtttg tcctgtgcag cttctggctt caccttcacc gactatacca tggactgggt 2040
ccgtcaggcc ccgggtaagg gcctggaatg ggttgcatg gttaatccta acagtggcgg 2100
ctctatctat aaccagcgt tcaagggcgg ttctactctg agtggtgaca gatctaataa 2160
cacattatac ctgcagatga acagcctgcg tgctgaggac actgccgtct attattgtgc 2220
tcgtaacctg ggacctctt tctactttga ctactggggt caaggaaccc tggtcaccgt 2280
ctcctcggcc tccaccaagg gcccatcggg cttccccctg gcacctcct ccaagagcac 2340
ctctgggggc acagcggccc tgggctgcct ggtcaaggac tacttccccg aaccgggtgac 2400
ggtgtcgtgg aactcaggcg ccctgaccag cggcgtgcac accttccccg ctgtcctaca 2460
gtcctcagga ctctactccc tcagcagcgt ggtgactgtg ccctctagca gcttgggac 2520
ccagacctac atctgcaacg tgaatcaca gccagcaac accaagggtg acaagaaagt 2580
tgagccaaa tctgtgaca aaactcacac atgccaccg tgcccagcac ctgaactcct 2640
ggggggaccg tcagtcttc tcttcccccc aaaaccgaag gacacctca tgatctccc 2700
gacctctgag gtcacatgcg tgggtgtgga cgtgagccac gaagaccctg aggtcaagtt 2760
caactgttac gtggacggcg tggaggtgca taatgccaag acaagccgc gggaggagca 2820
gtacaacagc agtaccggg tggtcagcgt cctcacgctc ctgcaccagg actggctgaa 2880
tggcaaggag tacaagtga aggtctcaa caagccctc ccagccccca togagaaaac 2940
catctccaaa gccaaagggc agccccgaga accacagggt tacacctgc cccatccc 3000
ggaagagatg accaagaacc aggtcagcct gacctgcctg gtcaaaggct tctatcccag 3060
cgacatcgcc gtggagtggg agagcaatgg gcagccggag aacaactaca agaccaogcc 3120
tcccgtgctg gactccgacg gctccttctt cctctacagc aagctcaccg tggacaagag 3180
caggtggcag caggggaacg tcttctcatg ctccgtgatg catgaggctc tgcacaacca 3240
ctacacgcag aagagcctct ccctgtctcc gggtaaata gaagcagcgc cctagagtcg 3300
acctgcagaa gcttcgatgg ccgccatggc ccaacttgtt tattgcagct tataatgggt 3360
acaaataaag caatagcatc acaaattca caaataaagc attttttca ctgcattcta 3420
gttgtggttt gtccaaactc atcaatgtat cttatcatgt ctggatcggg aattaattcg 3480
gcgcagcacc atggcctgaa ataacctctg aaagaggaac ttggttaggt accttctgag 3540
gcggaagaa ccagctgtgg aatgtgtgc agttagggtg tggaaagtcc ccaggctccc 3600
cagcaggcag aagtatgcaa agcatgcatc tcaattagtc agcaaccagg tgtggaaagt 3660
ccccaggctc ccagcaggc agaagtatgc aaagcatgca tctcaattag tcagcaacca 3720

tagtcccgcc	cctaactccg	cccatcccgc	ccctaactcc	gcccagttcc	gcccattctc	3780
cgccccatgg	ctgactaatt	ttttttat	atgcagaggc	cgaggccgcc	tcggcctctg	3840
agctattcca	gaagtagtga	ggaggctttt	ttggaggact	aggcttttgc	aaaaagctag	3900
cttatccggc	cggaacggt	gcattggaac	gcggattccc	cgtgcccaaga	gtcaggtaag	3960
taccgcctat	agagtctata	ggcccacccc	cttggcttcg	ttagaacgcg	gctacaatta	4020
atacataacc	ttttggatog	atcctactga	cactgacatc	cactttttct	ttttctccac	4080
aggtgtccac	tcccaggtcc	aactgcaoct	cggttcgcga	agctagcttg	ggctgcatcg	4140
attgaattcc	accatgggat	ggcatgtat	catccttttt	ctagtagcaa	ctgcaactgg	4200
agtacattca	gatatccaga	tgaccagtc	cccagactcc	ctgtccgcct	ctgtgggcga	4260
tagggtcacc	atcacctgca	aggccagtca	ggatgtgtct	attgggtctg	cctggtatca	4320
acagaaacca	ggaaaagctc	cgaaactact	gatttactcg	gcttcctacc	gatacactgg	4380
agtcccttct	cgcttctctg	gatccggttc	tgggacggat	ttcactctga	ccatcagcag	4440
tctgcagcca	gaagacttcg	caacttatta	ctgtcaacaa	tattatattt	atccttacac	4500
gtttgacag	ggtaccaagg	tggagatcaa	acgaactgtg	gctgcaccat	ctgtcttcat	4560
cttcccgcc	tctgatgagc	agttgaaatc	tggaactgct	tctgttgtgt	gcctgctgaa	4620
taacttctat	cccagagagg	ccaaagtaca	gtggaagggtg	gataacgccc	tccaatcggg	4680
taactcccag	gagagtgtca	cagagcagga	cagcaaggac	agcacctaca	gcctcagcag	4740
cacctgacg	ctgagcaaag	cagactacga	gaaacacaaa	gtctacgcct	gcgaagtcac	4800
ccatcagggc	ctgagctcgc	ccgtcacaaa	gagcttcaac	aggggagagt	gttaagcttc	4860
gatggccgcc	atggccaac	ttgtttattg	cagcttataa	tggttacaaa	taaagcaata	4920
gcatcacaaa	tttcacaaat	aaagcatttt	tttactgca	ttctagtgtg	ggtttgtcca	4980
aactcatcaa	tgtatcttat	catgtctgga	tcgggaatta	attcggcgca	gcaccatggc	5040
ctgaaataag	tttaaaccct	ctgaaagagg	aacttgggta	ggtaccgact	agtagcaagg	5100
tcgccagca	caagatcaat	attaacaatc	agtcactctc	ctttagcaat	aaaaggtga	5160
aaaattacat	tttaaaaatg	acaccataga	cgatgtatga	aaataatcta	cttgaaata	5220
aatctaggca	aagaagtgca	agactgttac	ccagaaaact	tacaaattgt	aatgagagg	5280
ttagtgaaga	tttaaatgaa	tgaagatcta	aataaactta	taaattgtga	gagaaattaa	5340
tgaatgtcta	agttaatgca	gaaacggaga	gacatactat	attcatgaac	taaaagactt	5400

aatattgtga aggtatactt tcttttcaca taaatttga gtcaatatgt tcaccccaaa 5460
aaagctgttt gttaacttgt caacctcatt tcaaaatgta tatagaaagc ccaaagacaa 5520
taacaaaaat attctttag aacaaaatgg gaaagaatgt tccactaaat atcaagattt 5580
agagcaaagc atgagatgtg tggggataga cagtgaggct gataaaatag agtagagctc 5640
agaaacagac ccattgatat atgtaagtga cctatgaaaa aaatatggca ttttacaatg 5700
ggaaaatgat gatctttttc ttttttagaa aaacagggaa atatatttat atgtaaaaaa 5760
taaaagggaa cccatatgtc ataccatata cacaaaaaaa ttccagtga ttataagtct 5820
aatggagaa ggcaaaactt taaatctttt agaaaataat atagaagcat gccatcatga 5880
cttcagtgtg gagaaaaatt tcttatgact caaagtccta accacaaaga aaagattggt 5940
aattagattg catgaatatt aagacttatt tttaaaatta aaaaaccatt aagaaaagtc 6000
aggccataga atgacagaaa atatttcaa caccocagta aagagaattg taatatgcag 6060
attataaaaa gaagtcttac aatcagtaa aaaataaaac tagacaaaaa tttgaacaga 6120
tgaaagagaa actctaaata atcattacac atgagaaact caatctcaga aatcagagaa 6180
ctatcattgc atatacacta aattagagaa atattaaaag gctaagtaac atctgtggca 6240
atattgatgg tatataacct tgatatgatg tgatgagaac agtactttac cccatgggct 6300
tcctcccaaa acccttacc cagtataaat catgacaaat atactttaa aaccattacc 6360
ctatatctaa ccagtactcc tcaaaactgt caaggtcatc aaaaataaga aaagtctgag 6420
gaactgtcaa aactaagagg aaccaagga gacatgagaa ttatatgtaa tgtggcattc 6480
tgaatgagat ccgagaacag aaaaagaaca gtagctaaaa aactaatgaa atataaataa 6540
agtttgaact ttagtttttt ttaaaaaga gtagcattaa cacggcaaag tcattttcat 6600
atitttcttg aacattaagt acaagtctat aattaaaaat tttttaaag tagtctggaa 6660
cattgccaga aacagaagta cagcagctat ctgtgctgtc gcctaactat ccatagctga 6720
ttggtctaaa atgagatata tcaacgctcc tccatgtttt ttgttttctt tttaaatgaa 6780
aaactttatt ttttaagagg agtttcaggt tcatagcaaa attgagagga aggtacattc 6840
aagctgagga agttttctc tattcctagt ttactgagag attgcatcat gaatgggtg 6900
taaattttgt caaatgcttt ttctgtgtct atcaatatga ccatgtgatt ttcttctta 6960
acctgttgat gggacaaatt acgttaattg attttcaaac gttgaaccac ccttacatat 7020
ctggaataaa ttctacttgg ttgtggtgta tattttttga tacattcttg gattcttttt 7080
gctaataatt tgttgaaaat gtttgtatct ttgttcatga gagatattgg tctgttgttt 7140

tcttttcttg taatgtcatt ttctagttcc ggtattaagg taatgctggc ctagttgaat 7200
gatttaggaa gtattccctc tgcttctgtc ttctgaggta ccgcgccgc cgcgctttt 7260
acaacgtcgt gactgggaaa accctggcgt taccoaactt aatogccttg cagcacatcc 7320
ccctttcgcc agctggcgta atagcgaaga ggcccgcacc gatcgccctt cccaacagtt 7380
gcgcagcctg aatggcgaat ggcgcctgat gcggtatfff ctcttacgc atctgtgcgg 7440
tatttcacac cgcatacgtc aaagcaacca tagtacgcgc cctgtagcgg cgcattaagc 7500
gcggcgggtg tgggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgcgc 7560
gctcctttcg ctttcttccc ttctttctc gccacgttcg ccggtttcc cgtcaagct 7620
ctaaatcggg ggctcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa 7680
aaacttgatt tgggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc 7740
cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaaacaaca 7800
ctcaacccta tctcgggcta ttcttttgat ttataaggga ttttgccgat ttccggcctat 7860
tggttaaaaa atgagctgat ttaacaaaa ttaacgcga attttaaca aatattaacg 7920
tttacaattt tatggtgcac tctcagtaca atctgctctg atgccgata gttaagccag 7980
ccccgacacc cgccaacacc cgctgacgcg ccctgacggg cttgtctgct ccggcctacc 8040
gcttacagac aagctgtgac cgtctccggg agctgcatgt gtcagagggt ttcaccgtca 8100
tcaccgaaac gcgcgagaga cgaaaggcc tcgtgatacg cctatffffa taggttaatg 8160
tcatgataat aatggtttct tagacgtcag gtggcacttt tcggggaaat gtgcgcggaa 8220
ccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 8280
cctgataaat gcttcaataa tattgaaaa ggaagagtat gagtattcaa catttccgtg 8340
tcgcccttat tccctttttt gcggcatttt gccttctgt ttttgctcac ccagaaacgc 8400
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 8460
atctcaacag cggtaagatc cttgagagtt ttcccccga agaacgtttt ccaatgatga 8520
gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc 8580
aactcggctg ccgcatacac tattctcaga atgacttggg tgagtactca ccagtcacag 8640
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgtgcc ataaccatga 8700
gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagetaaccg 8760
cttttttgca caacatgggg gatcatgtaa ctgccttga tcgttgggaa ccggagctga 8820

atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt 8880
tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 8940
ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt 9000
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 9060
ggccagatgg taagccctcc cgtatcgtag ttatotacac gacggggagt caggcaacta 9120
tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac 9180
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 9240
aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt 9300
tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt 9360
ttttctgcy cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt 9420
gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 9480
agatacaaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 9540
tagcaccgcc tacatactc gctctgctaa tctgttacc agtggctgct gccagtggcg 9600
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggg 9660
cgggctgaac ggggggttcg tgcacacagc ccagcttggg gcgaacgacc tacaccgaac 9720
tgagatacct acagcgtgag ctatgagaaa gcgccacgct tccgaaggg agaaggcgg 9780
acaggtatcc ggtaagcggc agggctcgaa caggagagcg cacgaggag cttccagggg 9840
gaaacgcctg gtatctttat agtctgtcgc ggtttcgcca cctctgactt gagcgtcgat 9900
ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt 9960
tacggttctt ggcttttgc tggccttttg ctcacatggt ctttctgcy ttatcccctg 10020
attctgtgga taaccgtatt accgcctttg agtgagctga tacogctcgc cgcagccgaa 10080
cgaccgagcg cagcagatca gtgagcaggg aagcgggaaga gcccgcgggc aaggctcgcca 10140
cgcacaagat caatattaac aatcagtcac ctctctttag caataaaaag gtgaaaaatt 10200
acattttaa aatgacacca tagacgatgt atgaaaataa tctacttggg aataaatcta 10260
ggcaaagaag tgcaagactg ttaccagaa aacttacaaa ttgtaatga gaggttagtg 10320
aagatttaa tgaatgaaga tctaaataa cttataaatt gtgagagaaa ttaatgaatg 10380
tctaagttaa tgcagaaaac gagagacata ctatattcat gaactaaaag acttaatatt 10440
gtgaaggtat actttctttt cacataaatt tgtagtcaat atgttcacc caaaaagct 10500
gtttgtaac ttgtcaacct catttcaaaa tgtatataga aagcccaag acaataacaa 10560

aaatattctt gtagaacaaa atgggaaaga atgttcact aaatatcaag atttagagca 10620
aagcatgaga tgtgtgggga tagacagtga ggctgataaa atagagtaga gctcagaaac 10680
agaccattg atatatgtaa gtgacctatg aaaaaatat ggcattttac aatgggaaaa 10740
tgatgatctt tttctttttt agaaaaacag ggaaatatat ttatatgtaa aaaataaaag 10800
ggaaccata tgcatacca tacacacaaa aaaattccag tgaattataa gtctaaatgg 10860
agaaggcaaa actttaaatc ttttagaaaa taatatagaa gcatgccatc atgacttcag 10920
tgtagagaaa aatttcttat gactcaaagt cctaaccaca aagaaaagat tgtaattag 10980
attgcatgaa tattaagact tatttttaaa attaaaaaac cattaagaaa agtcaggcca 11040
tagaatgaca gaaaatattt gcaacacccc agtaaagaga attgtaatat gcagattata 11100
aaaagaagtc ttacaaatca gtaaaaaata aaactagaca aaaatttgaa cagatgaaag 11160
agaaactcta aataatcatt acacatgaga aactcaatct cagaaatcag agaactatca 11220
ttgcatatac actaaattag agaaatatta aaaggctaag taacatctgt ggcaatattg 11280
atggtatata accttgatat gatgtgatga gaacagtact ttaccccatg ggcttcctcc 11340
ccaaaccctt accccagtat aaatcatgac aaatatactt taaaaccat taccctatat 11400
ctaaccagta ctctcaaaa ctgtcaaggt catcaaaaat aagaaaagtc tgaggaactg 11460
tcaaaactaa gaggaacca aggagacatg agaattatat gtaatgtggc attctgaatg 11520
agatcccaga acagaaaaag aacagtagct aaaaaactaa tgaaatataa ataaagttg 11580
aactttagtt ttttttaaaa aagagtagca ttaacacggc aaagtcattt tcatattttt 11640
cttgaacatt aagtacaagt ctataattaa aaatttttta aatgtagtct ggaacattgc 11700
cagaaacaga agtacagcag ctatctgtgc tgtcgctaa ctatccatag ctgattggtc 11760
taaaatgaga tacatcaacg ctctccatg tttttgttt tctttttaa tgaaaaactt 11820
tattttttaa gaggagtttc aggttcatag caaaattgag aggaaggtac attcaagctg 11880
aggaagtttt cctctattcc tagtttactg agagattgca tcatgaatgg gtgttaaatt 11940
ttgtcaaagc cttttctgt gtctatcaat atgaccatgt gattttctc tttaacctgt 12000
tgatgggaca aattacgtta attgattttc aaacgttgaa ccacccttac atatctggaa 12060
taaattctac ttggttggtg tgtatatttt ttgatacatt cttggattct ttttgctaat 12120
attttgttga aaatgtttgt atctttgttc atgagagata ttggtctgtt gttttctttt 12180
cttghtaatgt cattttctag ttccggtatt aaggtaatgc tggcctagtt gaatgattta 12240

ggaagtattc cctctgcttc tgtcttctga agcggaagag cgccaatac gcaaaccgcc 12300
tctccccgcg cgttgccga ttcattaatg cagctggcac gacaggttc ccgactggaa 12360
agcgggcagt gagcgcaacg caattaatgt gagttagctc actcattagg caccccaggc 12420
tttaccttt atgcttccg ctcgtatgtt gtgtggaatt gtgagcggat aacaatttca 12480
cacaggaac agctatgaca tgattacgaa ttaa 12514

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT OR DRAWING
- BLURRED OR ILLEGIBLE TEXT OR DRAWING
- SKEWED/SLANTED IMAGES
- COLOR OR BLACK AND WHITE PHOTOGRAPHS
- GRAY SCALE DOCUMENTS
- LINES OR MARKS ON ORIGINAL DOCUMENT
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.